

OM of: US-09-528-682-1 to: Issued_Patents_NA:* out_format : pfs
Date: Jun 18, 2002 10:50 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command Line Parameters:

```
-MODEL=frame+g2n.model -DEV=xlp  
-O=/cgn2_1/USPTO_spool/US09528682/runat_18062002_082606_8822/app_query.fasta_1.299  
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olip2n.rni  
-GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPL=0.000  
-LOAPEXT=0.000 -GAPEXT=4.500 -GAPEXT=0.050 -XGAPOP=60.000  
-XGAPEXT=60.000 -XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=60.000  
-XGAPEXT=60.000 -DELOP=6.000 -DELEX=7.000 -START=1  
-MATRIX=oligo -TRANS=human40.cdi -LIST=1000 -DOCALLIGN=200  
-THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-ORNext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-SER=US09528682 -RCGN1_1.80 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1
```

Search Information block:

Query: US-09-528-682-1
Query Length: 240
Database: Issued_Patents_NA
Database sequences: 383533
Database length: 122816752
Search time (sec): 77.150000

WARN: XGAPOP and XGAPEXT must be equal. Assuming XGAPOP=XGAPEXT=60.000
WARN: XGAPEXT and XGAPEXT must be equal. Assuming XGAPEXT=XGAPEXT=60.000

Score_List:

Sequence	Strd Orig	ZScore	EScore Len	Documentation	711	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-823-120-5 +		86.00	1606.33	8.7e-82	711	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-435-605A-7 +		26.00	472.28	1.3e-18	576	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-435-605A-5 +		26.00	472.21	1.3e-18	582	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-823-120-7 +		26.00	470.61	1.6e-18	723	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-435-605A-1 +		26.00	470.11	1.7e-18	774	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-449-045C-1 +		26.00	470.08	1.7e-18	777	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-296-848A-1 +		15.00	282.82	4.6e-08	45	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-435-605A-9 +		12.00	217.81	0.0002	138	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-449-045C-23 +		11.00	206.34	0.0008	50	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-435-605A-35 +		11.00	206.34	0.0008	50	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-435-605A-36 +		11.00	206.34	0.0008	50	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-629-629-14 +		10.00	188.69	0.0080	42	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-435-605A-35 +		8.00	153.82	0.7028	28	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-403-58A-5 +		8.00	152.84	0.7970	32	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-367-115-5 +		8.00	152.84	0.7970	32	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-367-115-5 +		8.00	152.84	0.7970	32	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-367-115-5 +		8.00	152.84	0.7970	32	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-296-848A-4 +		8.00	150.33	1.10	45	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-296-848A-1 +		8.00	150.33	1.10	45	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-07-932-379A-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-07-932-379A-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-07-932-379A-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5A.COMB.seq:US-08-379-295-12 +		7.00	136.66	6.34	22	1	9
/cgn2_6/pdata/2/lna/5B.COMB.seq:US-08-379-295-11 +		7.00					

/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-483-577A-3	-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-295-802-140	+	6.00	113.34	126.23	40
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-837-438-3	-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-868-986-67	+	6.00	113.34	126.23	40
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-637-654-3	-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-486-837A-140	+	6.00	113.34	126.23	40
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-487-890A-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-375-992A-140	+	6.00	113.34	126.23	40
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-478-435-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-185-949-98	+	6.00	112.99	133.17	42
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-337-483-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-449-045-35	+	6.00	111.71	155.76	50
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-478-373-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-435-605-37	+	6.00	111.71	155.76	50
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-474-671-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-330-867A-41	+	6.00	111.71	155.76	50
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-883-577A-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-956-260-41	+	6.00	111.71	155.76	50
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-897-438-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-448-570A-48	+	6.00	109.67	202.32	66
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-637-654-2	-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-188-877A-38	+	6.00	109.67	202.32	66
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-390-721-1	-	7.00	96.74	1.1e+03	5045	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-918-148-38	+	6.00	108.44	236.81	78
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-487-890A-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-680-326-44	-	6.00	108.44	236.81	78
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-478-435-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-172-711-61	-	6.00	106.99	285.14	95
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-337-483-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-178-015-5	-	6.00	106.99	285.14	95
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-478-373-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-337-668-48	+	6.00	103.69	435.70	149
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-474-671-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-484-570A-48	+	6.00	103.64	435.46	150
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-483-577A-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-888-077A-38	+	6.00	103.34	443.96	152
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-897-438-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-136-277-10	-	6.00	103.54	443.96	152
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-637-654-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-479-403-10	+	6.00	103.54	443.96	152
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-837-654-4	-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-835-734-10	-	6.00	103.35	454.96	153
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-478-435-105	-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-457-904A-28	-	6.00	101.09	607.38	122
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-637-654-105	-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-457-904A-28	-	6.00	101.09	607.38	122
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-977-405A-1	+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-456-701A-18	-	6.00	101.09	607.38	122
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-977-871A-1	+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-456-701A-28	-	6.00	101.09	607.38	122
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-925-951-1	+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-456-701A-36	+	6.00	101.06	610.08	213
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-120-663-1	-	7.00	96.08	1.2e+03	5151	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-435-684A-36	+	6.00	101.06	610.08	213
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-450-332-1	-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-934-877A-36	+	6.00	101.06	610.08	213
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-637-640-1	-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-934-877A-36	+	6.00	101.06	610.08	213
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-004-406C-1	-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-871-678C-41	+	6.00	101.06	610.08	213
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-871-678C-41	+	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100.69	639.72	224
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-367-841A-43	-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-700-575-19	-	6.00	100		

cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-4702-344-11	6.00	97.67	942.54	338	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-4873-031-14	6.00	97.52	960.92	345	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-473-034-14	6.00	97.52	960.92	345	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-0670-756-61	6.00	97.29	989.76	353	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-136-277-21	6.00	96.52	1.1e+03	395	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-479-403-21	6.00	96.52	1.1e+03	395	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-835-734-21	6.00	96.47	1.1e+03	398	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-07-846-992-2	6.00	96.47	1.1e+03	398	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-469-553-2	6.00	96.41	1.1e+03	398	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-896-164-83	6.00	96.41	1.1e+03	401	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-318-193-88	6.00	96.20	1.1e+03	412	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-227-357-71	6.00	96.20	1.1e+03	413	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5496924-56	6.00	96.20	1.1e+03	413	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-1965-388-1	6.00	96.02	1.2e+03	423	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-060-756-60	6.00	95.77	1.2e+03	438	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-09-323-873A-11	6.00	95.60	1.2e+03	448	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-387-942C-38	6.00	95.52	1.2e+03	453	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-680-395-8	6.00	95.49	1.2e+03	455	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-09-0670-756-60	6.00	95.42	1.3e+03	459	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-184-009-139	6.00	95.37	1.3e+03	462	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-458-356-139	6.00	95.37	1.3e+03	462	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-012-366-1	6.00	95.37	1.3e+03	462	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-460-736-139	6.00	95.37	1.3e+03	462	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-160-567-10	6.00	95.25	1.3e+03	470	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-031-143B-73	6.00	94.95	1.3e+03	470	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-031-143B-73	6.00	94.94	1.3e+03	490	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-060-756-587	6.00	94.91	1.3e+03	492	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5219739-21	6.00	94.82	1.4e+03	498	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-928-692-56	6.00	94.78	1.4e+03	501	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-040-833A-134	6.00	94.75	1.4e+03	503	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-08-545-809A-21	6.00	94.55	1.4e+03	519	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-487-031-11	6.00	94.35	1.4e+03	531	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-473-034-11	6.00	94.35	1.4e+03	531	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-468-946-3	6.00	93.84	1.5e+03	569	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-348-572-6	6.00	93.84	1.5e+03	569	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-041-090B-6	6.00	93.84	1.5e+03	569	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-07-938-188-1	6.00	93.80	1.5e+03	572	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-306-062-1	6.00	93.80	1.5e+03	572	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-190-560-3	6.00	93.72	1.6e+03	579	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-469-277-3	6.00	93.72	1.6e+03	579	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5182210-11	6.00	93.56	1.6e+03	591	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-974-546-4	6.00	93.56	1.6e+03	593	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-974-546-4	6.00	93.56	1.6e+03	593	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-08-852-807-18	6.00	93.48	1.6e+03	598	1	5
cgnd_6/p/ptodate/2/1lna/5b/COMB.seq	US-09-221-298-9	6.00	93.40	1.6e+03	604	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5460029-2	6.00	93.40	1.6e+03	604	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5460029-2	6.00	93.40	1.6e+03	604	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-5460029-2	6.00	93.40	1.6e+03	604	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2/1lna/6b/COMB.seq	US-09-385-982-291	6.00	93.34	1.6e+03	609	1	5
cgnd_6/p/ptodate/2							

/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-925-091-6-	6.00	89.16	2.8e+03	1076	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-177-776-43+	6.00	87.22	3.56e+03	1407
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-077-675A-6+	6.00	89.08	2.8e+03	1088	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-151-574-33+	6.00	87.21	3.56e+03	1400
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-022-411-1-	6.00	89.06	2.8e+03	1091	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-166-424-19+	6.00	87.21	3.56e+03	1404
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-716-161A-3-	6.00	89.03	2.8e+03	1096	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-693-408-11+	6.00	87.21	3.56e+03	1404
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-944-604-19-	6.00	89.01	2.9e+03	1099	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-419-745-33+	6.00	87.21	3.56e+03	1400
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-716-161A-10-	6.00	88.92	2.9e+03	1112	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-923-510-33+	6.00	87.19	3.56e+03	1407
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-077-675A-9+	6.00	88.86	2.9e+03	1122	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-598-876-11+	6.00	87.18	3.56e+03	1400
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-767-096-1+	6.00	88.62	3.0e+03	1158	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-605-830-69+	6.00	87.18	3.56e+03	1400
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-480-203-1+	6.00	88.62	3.0e+03	1158	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-368-704C-50+	6.00	86.94	3.7e+03	1403
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-227-357-41+	6.00	88.54	3.0e+03	1171	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-934-613-60-	6.00	86.92	3.7e+03	1403
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-107-676-2-	6.00	88.50	3.1e+03	1178	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-09-234-613-60-	6.00	86.91	3.7e+03	1462
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-890-719-10-	6.00	88.39	3.1e+03	1195	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-775-271-3+	6.00	86.91	3.7e+03	1462
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-018-584A-37-	6.00	88.36	3.1e+03	1200	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-09-004-838-126-	6.00	86.90	3.7e+03	1468
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-107-676-2-	6.00	88.30	3.1e+03	1211	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-330-1218-1-	6.00	86.81	3.8e+03	1481
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-829-287-3-	6.00	88.23	3.2e+03	1221	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-270-076A-2+	6.00	86.81	3.8e+03	1489
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-920-719-2+	6.00	88.18	3.2e+03	1230	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-828-741B-3+	6.00	86.77	3.8e+03	1490
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-980-719-9+	6.00	88.18	3.2e+03	1230	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-09-160-567-3+	6.00	86.77	3.8e+03	1491
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-240-372-1+	6.00	88.04	3.2e+03	1237	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-833-876-1-	6.00	86.77	3.8e+03	1491
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-591-629-1+	6.00	88.04	3.2e+03	1253	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-483-054-1-	6.00	86.77	3.8e+03	1491
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-642-274D-49-	6.00	88.00	3.2e+03	1254	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-632-470-51-	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-952-014C-49-	6.00	88.00	3.3e+03	1261	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-349-025-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-745-742-16+	6.00	87.99	3.3e+03	1261	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-668-950A-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-989-386-4-	6.00	87.99	3.3e+03	1261	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-656-568-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-858-207A-162+	6.00	87.95	3.3e+03	1269	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-200-673-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-878-588-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-740-223A-17-	6.00	86.72	3.8e+03	1501
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-483-232-23+	6.00	86.56	3.9e+03	1533
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-961-083-171+	6.00	87.88	3.3e+03	1282	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-485-938A-23+	6.00	86.56	3.9e+03	1533
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-040-127-1+	6.00	87.76	3.4e+03	1302	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.56	3.9e+03	1534
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-041-011-1+	6.00	87.76	3.4e+03	1302	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-814-877-1+	6.00	86.56	3.9e+03	1534
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-604-17-	6.00	87.60	3.4e+03	1331	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-07-494-925-11-	6.00	86.53	3.9e+03	1540
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-922-957-2-	6.00	87.58	3.4e+03	1334	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-09-100-546-23+	6.00	86.56	3.9e+03	1533
/cgn2_6/p/ptodata/2/1na/6B.COMB	seq:US-08-404-127-3+	6.00	87.58	3.4e+03	1335	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-837-593-1+	6.00	86.51	3.9e+03	1544
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-041-011-3+	6.00	87.58	3.4e+03	1335	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-439-04392-1+	6.00	86.47	4.0e+03	1568
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-358-287B-41-	6.00	87.50	3.5e+03	1350	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-810-712-3-	6.00	86.40	4.0e+03	1568
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-368-704C-41-	6.00	87.50	3.5e+03	1350	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-09-173-581-12+	6.00	86.37	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-908-403-63-1-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-09-420-915-2+	6.00	86.37	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-083-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-09-420-915-2+	6.00	86.37	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-286-304-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5A.COMB	seq:US-08-944-816-1+	6.00	86.32	4.0e+03	1574
/cgn2_6/p/ptodata/2/1na/6A.COMB	seq:US-08-443-129-2-	6.00	87.49	3.5e+03	1352	/cgn2_6/p/ptodata/2/1na/5B.COMB	seq:US-08-944-816-1+	6			

/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-591-468-6 +	6.00	85.96	4.2e+03	1665	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-390-888A-1 -	6.00	84.77	4.9e+03	1958
/cgn2_6/pdata/2/1na/6A.COMB.seq:PCP-US94-0640-6 +	6.00	85.94	4.2e+03	1669	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-803-636A-1 -	6.00	84.75	4.9e+03	1962
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-964-308-3 +	6.00	85.94	4.2e+03	1669	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-106-761-3 +	6.00	84.73	5.0e+03	1969
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-063-138-3 -	6.00	85.94	4.2e+03	1669	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-188-277B-3 +	6.00	84.73	5.0e+03	1969
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-003-159-1 +	6.00	85.93	4.2e+03	1672	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-429-844-79 +	6.00	84.73	5.0e+03	1969
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-360-673-1 +	6.00	85.87	4.3e+03	1685	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-596-867C-1 -	6.00	84.72	5.0e+03	1970
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-648-657-14 +	6.00	85.86	4.3e+03	1686	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-450-933A-3 +	6.00	84.69	5.0e+03	1979
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-240-012-12 -	6.00	85.84	4.3e+03	1692	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-465-669-3 +	6.00	84.69	5.0e+03	1979
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-099-677A-4 -	6.00	85.80	4.3e+03	1700	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-257-963B-11 +	6.00	84.65	5.0e+03	1988
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-261-471-4 -	6.00	85.80	4.3e+03	1708	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:PCP-US95-07201-11 +	6.00	84.65	5.0e+03	1988
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-108-020-5 +	6.00	85.75	4.3e+03	1713	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-798-776-1 +	6.00	84.56	5.1e+03	2014
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-379-482A-2 +	6.00	85.64	4.4e+03	1718	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-228-819A-1 +	6.00	84.56	5.1e+03	2014
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-173-508-5 -	6.00	85.48	4.5e+03	1777	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-980-514-2 +	6.00	84.52	5.1e+03	2018
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-263-310-5 -	6.00	85.48	4.5e+03	1777	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-937-540-1 -	6.00	84.52	5.1e+03	2025
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-951-742-5 -	6.00	85.47	4.5e+03	1780	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-937-540-1 -	6.00	84.49	5.1e+03	2034
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-933-821-5 -	6.00	85.47	4.5e+03	1780	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-294-923-2 +	6.00	84.49	5.1e+03	2034
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-960-507-5 -	6.00	85.47	4.5e+03	1780	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-657-661-3 +	6.00	84.40	5.2e+03	2059
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-136-828-5 -	6.00	85.47	4.5e+03	1780	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:PCP-US94-07233-3 +	6.00	84.40	5.2e+03	2059
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-381-603-3 -	6.00	85.46	4.5e+03	1782	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-750-391-6 -	6.00	84.35	5.2e+03	2072
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-924-376-3 -	6.00	85.46	4.5e+03	1782	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-367-841A-11 +	6.00	84.34	5.2e+03	2075
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-685-212-3 -	6.00	85.46	4.5e+03	1782	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-750-391-6 -	6.00	84.33	5.2e+03	2078
/cgn2_6/pdata/2/1na/PCPUS.COMB.seq:PCP-US94-02414-3 -	6.00	85.46	4.5e+03	1782	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-299-843A-65 +	6.00	84.30	5.2e+03	2085
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-379-482A-2 +	6.00	85.46	4.5e+03	1782	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-445-520B-8 -	6.00	84.29	5.2e+03	2085
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-920-430-19 +	6.00	85.44	4.5e+03	1786	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-451-948B-7 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-065-299-9 +	6.00	85.44	4.5e+03	1786	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-445-948B-7 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-263-047-1 +	6.00	85.44	4.5e+03	1786	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-445-948B-7 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-251-064-19 +	6.00	85.44	4.5e+03	1786	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-311-703A-7 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-379-482A-2 +	6.00	85.41	4.5e+03	1794	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-189-543-7 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-785-396-11 +	6.00	85.41	4.5e+03	1794	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-445-520B-8 -	6.00	84.29	5.2e+03	2090
/cgn2_6/pdata/2/1na/PCPUS.COMB.seq:PCP-US95-14418-3 +	6.00	85.41	4.5e+03	1794	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-233-864A-10 -	6.00	84.27	5.2e+03	2090
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-733-828-1 +	6.00	85.38	4.6e+03	1800	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-367-841A-11 +	6.00	84.25	5.3e+03	2102
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-591-468-6 +	6.00	85.34	4.6e+03	1810	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-063-237-3 -	6.00	84.23	5.3e+03	2105
/cgn2_6/pdata/2/1na/5B.COMB.seq:PCP-US94-06430-10 +	6.00	85.34	4.6e+03	1810	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-07-955-041-3 -	6.00	84.23	5.3e+03	2105
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-988-876-2 +	6.00	85.27	4.6e+03	1828	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-427-482-3 -	6.00	84.23	5.3e+03	2105
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-009-075-3 -	6.00	85.26	4.6e+03	1830	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-481-069-3 -	6.00	84.23	5.3e+03	2105
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-338-671-3 +	6.00	85.24	4.6e+03	1835	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-481-069-3 -	6.00	84.23	5.3e+03	2105
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-665-926-7 +	6.00	85.19	4.7e+03	1849	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-032-324-5 +	6.00	84.09	5.4e+03	2146
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-740-223A-9 +	6.00	85.19	4.7e+03	1849	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-382-326-15 -	6.00	84.05	5.4e+03	2146
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-976-259-95 +	6.00	85.12	4.7e+03	1866	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-395-115-15 -	6.00	84.05	5.4e+03	2160
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-173-581-13 -	6.00	85.12	4.7e+03	1866	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-436-265-13 -	6.00	84.02	5.4e+03	2167
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-510-801-1 +	6.00	85.12	4.7e+03	1867	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-368-704C-52 +	6.00	84.01	5.4e+03	2170
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-248-355-1 +	6.00	85.12	4.7e+03	1867	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-173-639-1 -	6.00	84.00	5.4e+03	2170
/cgn2_6/pdata/2/1na/5B.COMB.seq:PCP-US95-06683-1 +	6.00	85.08	4.7e+03	1875	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-258-287B-54 +	6.00	83.99	5.4e+03	2174
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-965-139-14 +	6.00	85.07	4.7e+03	1878	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-357-251-9 -	6.00	83.97	5.4e+03	2182
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-995-659-14 +	6.00	85.07	4.7e+03	1878	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-845-528-8 -	6.00	83.94	5.4e+03	2191
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-215-649A-14 +	6.00	85.07	4.7e+03	1878	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-989-362-1 +	6.00	83.94	5.4e+03	2191
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-431-080-15 +	6.00	85.06	4.8e+03	1882	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-033-392-1 +	6.00	83.94	5.4e+03	2192
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-931-080-29 -	6.00	85.06	4.8e+03	1882	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-504-511A-1 +	6.00	83.90	5.4e+03	2204
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-938-534-29 -	6.00	85.06	4.8e+03	1882	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-07-955-800-3 -	6.00	83.85	5.4e+03	2219
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-938-534-29 -	6.00	85.03	4.8e+03	1890	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:PCP-US91-00399-7 -	6.00	83.85	5.4e+03	2219
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-503-541B-11 +	6.00	84.94	4.8e+03	1912	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-311-81-3 +	6.00	83.78	5.4e+03	2223
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-868-435-11 +	6.00	84.94	4.8e+03	1912	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-890-853-3 +	6.00	83.78	5.4e+03	2223
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-744-231-11 +	6.00	84.88	4.9e+03	1927	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-033-476-3 +	6.00	83.78	5.4e+03	2238
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-336-583-1 +	6.00	84.88	4.9e+03	1927	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-099-125A-3 +	6.00	83.78	5.4e+03	2238
/cgn2_6/pdata/2/1na/PCPUS.COMB.seq:PCP-US95-13795-1 -	6.00	84.83	4.9e+03	1940	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-463-379B-34 -	6.00	83.75	5.4e+03	2249
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-233-200-28 +	6.00	84.83	4.9e+03	1941	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-463-379B-34 -	6.00	83.75	5.4e+03	2249
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-09-233-197-28 +	6.00	84.83	4.9e+03	1941	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-463-379B-34 -	6.00	83.75	5.4e+03	2249
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-09-233-201-28 +	6.00	84.83	4.9e+03	1941	1	/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-463-379B-34 -	6.00	83.75	5.4e+03	2249
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-540-804-1 -	6.00	84.80	4.9e+03	1949	1	/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-465-585C-34 -	6.00	83.75	5.4e+03	2249
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-218-265-1 -	6.00	84.80	4.9e+03	1949	1					
/cgn2_6/pdata/2/1na/6A.COMB.seq:US-08-521-872-1 -	6.00	84.80	4.9e+03	1949	1					
/cgn2_6/pdata/2/1na/5B.COMB.seq:US-08-590-399-1 -	6.00	84.80	4.9e+03	1949	1					


```

/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-652-446-34- 6.00 83.75 5.6e+03 2249 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-910-820-7+ 6.00 83.74 5.6e+03 2251 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-322-742-2+ 6.00 83.71 5.6e+03 2261 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-276-400-3+ 6.00 83.71 5.6e+03 2262 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-448-076-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-176-126B-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-669-435-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:PCR-US94-14431A-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-444-005-14+ 6.00 83.69 5.7e+03 2268 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-197-360-1+ 6.00 83.67 5.7e+03 2273 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-168-629-1+ 6.00 83.67 5.7e+03 2273 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-810-131A-1+ 6.00 83.60 5.7e+03 2295 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-842-842-6- 6.00 83.60 5.7e+03 2295 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-09-052-521C-1- 6.00 83.52 5.8e+03 2319 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-286-767-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-090-526-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-486-715-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-486-719-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-271-354-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-476-100-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-565-861-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-475-749-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:PCR-US94-07658-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-975-762-44- 6.00 83.51 5.8e+03 2322 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-295-028-44- 6.00 83.51 5.8e+03 2322 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-106-582-44- 6.00 83.50 5.8e+03 2322 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-868-435-1+ 6.00 83.50 5.8e+03 2327 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-744-231-1+ 6.00 83.49 5.8e+03 2327 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-811-481-34- 6.00 83.49 5.8e+03 2330 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-557-128-3- 6.00 83.49 5.8e+03 2330 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-242-690A-34- 6.00 83.46 5.8e+03 2330 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-09-022-983-4+ 6.00 83.46 5.8e+03 2335 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-07-821-716-3- 6.00 83.39 5.9e+03 2363 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-07-923-724-7+ 6.00 83.39 5.9e+03 2363 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-609-426A-7+ 6.00 83.39 5.9e+03 2363 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-981-659-1- 6.00 83.38 5.9e+03 2369 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-883-534-4+ 6.00 83.37 5.9e+03 2369 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-09-204-764-4+ 6.00 83.37 5.9e+03 2379 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-374-652C-1+ 6.00 83.34 5.9e+03 2382 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-468-558-4- 6.00 83.33 5.9e+03 2382 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-676-444-4- 6.00 83.29 6.0e+03 2394 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-319-989-7- 6.00 83.26 6.0e+03 2404 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-311-023-7- 6.00 83.19 6.0e+03 2427 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-490-099-1- 6.00 83.15 6.1e+03 2439 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-624-693A-18+ 6.00 83.13 6.1e+03 2445 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-858-207A-45+ 6.00 83.13 6.1e+03 2453 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-07-863-169A-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-429-964-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-07-935-087-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:PCR-US93-08062-4+ 6.00 83.07 6.1e+03 2464 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-490-692-3+ 6.00 83.03 6.2e+03 2481 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-630-118A-1+ 6.00 83.03 6.2e+03 2481 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-838-339-1+ 6.00 83.03 6.2e+03 2481 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-235-839-1+ 6.00 83.02 6.2e+03 2485 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-258-287B-52+ 6.00 82.99 6.2e+03 2492 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-258-287B-3+ 6.00 82.98 6.2e+03 2498 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-561-825-23- 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-073-384C-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-254-359A-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-483-043-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-481-238-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-471-066B-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-484-956-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-599-491-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-756-386-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-823-516-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-08-682-853A-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-759-038-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-758-316-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/prodata/2/lna/5B.COMB.seq:US-09-350-309-3+ 6.00 82.96 6.2e+03 2504 1

```


[illegible]

/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-975-762-42 - 6.00 79.92 9.1e+03 3786
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-09-295-028-42 - 6.00 79.92 9.1e+03 3786
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-09-106-582-42 - 6.00 79.92 9.1e+03 3786
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-07-872-644-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-297-494-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-297-510-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-454-455-5 + 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-479-532-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-455-526-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-455-525-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-09-139-491-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-09-139-491-42 - 6.00 79.92 9.1e+03 3789
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-936-165A-196 - 6.00 79.92 9.1e+03 3815
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-823-120-5

seq_name: /cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-823-120-5
seq_documentation_block:
; Sequence 5, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappunli, Rino
; APPLICANT: Pizze, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, their Preparation and
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,003
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..711
; US-08-823-120-5

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-823-120-5 ..

Align seq 1/1 to: US-08-823-120-5 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGlnGluValSerAlaIleuGlyIyl 119
295 GTATACAGCCCTCACCCATATGAACAGAGGTTTCTGCGTTAGGTGGAAT 344
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAlaSpHegIValIleA 136
345 ACCATATTTCACAGATATATGATGATGCTGTTAATTGCGTGAATTG 394
136 spGluArgLeuHisArgAsnArgGluTyrArgAspArgIylTyrArgAsn 152
395 ATGACACATTATACATCCGTACACGGGAATATATAGACACCGGTATTACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyHeProP 169
445 CTGATATATACCTCCGGCAGAGATGGTTACAGATTAGCAGGTTCCACC 494
169 CAAPHISGlnAlaTPrArgGluGluProTPrIleHisAlaProGln 186
495 GATCACCAACGCTTGGAGAGAACCCCTGATTTCATCATGCACCAAG 544
186 IycysGly 188
545 GTTGTGGA 552

seq_name: /cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-435-605A-7

seq_documentation_block:
; Sequence 7, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-435-605A-7

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-7 ..

Align seg 1/1 to: US-08-435-605A-7 from: 1 to: 576

29 GUUyrpheaSPARGlyThrGlnmetasnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTGTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-435-605A-5

seq_documentation_block:

Sequence 5, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
TITLE OF INVENTION: SUBUNIT ANALOGS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 582 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-605A-5

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-5 ..

Align seg 1/1 to: US-08-435-605A-5 from: 1 to: 582

29 GUUyrpheaSPARGlyThrGlnmetasnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTGTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-823-120-7

seq_documentation_block:

Sequence 7, Application US/08823120
Patent No. 6149919
GENERAL INFORMATION:
APPLICANT: Domenighini, Mario
APPLICANT: Rappuoli, Rino
APPLICANT: Pizza, Mariagrazia
TITLE OF INVENTION: Immunogenic Detoxified Mutants of
TITLE OF INVENTION: Cholera Toxin and of the Toxin B, Their Preparation and
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,120
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,003
FILING DATE: 11-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..723
US-08-823-120-7

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-823-120-7 ..

Align seg 1/1 to: US-08-823-120-7 from: 1 to: 723

29 GUUyrpheaSPARGlyThrGlnmetasnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTGTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-435-605A-1

seq_documentation_block:

Sequence 1, Application US/08435605A


```
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-435-605A-1

alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-435-605A-1 ..
  Align seg 1/1 to: US-08-435-605A-1 from: 1 to: 774

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
  |||||||
139 GAGTACTTGTACCGAGGTACTCAATGAATATCAACCTTATGATCATGTC 188
45 aargglythrGlnThrGlyPheValArg 54
  |||||||
189 AAGAGGAACCTCAGACGGGATTGTAGG 216

seq_name: /cgn2.6/ptodata/2/ina/5A.COMB.seq:US-08-449-045C-1

seq_documentation_block:
; Sequence 1, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..777
; US-08-449-045C-1
```

```
alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-449-045C-1 ..
  Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777
```

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
  |||||||
139 GAGTACTTGTACCGAGGTACTCAATGAATATCAACCTTATGATCATGTC 188
45 aargglythrGlnThrGlyPheValArg 54
  |||||||
189 AAGAGGAACCTCAGACGGGATTGTAGG 216
```

seq_name: /cgn2.6/ptodata/2/ina/backfiles1.seq:5223610-1

```
seq_documentation_block:
; Patent No. 5223610
; APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
; TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
; HORMONE PROMOTER
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,852
; FILING DATE: 18-MAY-1990
; SEQ ID NO: 1:
; LENGTH: 2020
; 5223610-1
```

```
alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:
 US-09-528-682-1 x 5223610-1 ..

Align seg 1/1 to: 5223610-1 from: 1 to: 2020

29 GltutyrheasparglythrGlnMetasnIleasnLeutyraPhsAI 45
|||||
654 GAGTACTTTGACCGAGTACTCAATGATATCACTTATATGATC 703
45 aargGlythrGlnThrGlypheValarg 54
|||||
704 AAGAGCACTCAGACGGGATTTGTAGG 731

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-296-848A-1

seq_documentation_block:

Sequence 1, Application US/08296848A
Patent No. 6019982

GENERAL INFORMATION:

APPLICANT: Clements, John D.

APPLICANT: Dickinson, Bonny L.

TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A

TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,848A

FILING DATE: 26-AUG-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5113-046

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA

US-08-296-848A-1

alignment_scores:

Quality: 15.00 Length: 15

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-296-848A-1 ..

Align seg 1/1 to: US-08-296-848A-1 from: 1 to: 45

186 GtYcsgIyasnSerArgThrIleThrGlyAspThrCysasn 200
|||||
1 GGTGTGGAATTCATCAAGACAAATTAACGATATCTGTAT 45

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-9

seq_documentation_block:

Sequence 9, Application US/08435605A

Patent No. 5874287

GENERAL INFORMATION:

APPLICANT: Burnette, W. Neal

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

TITLE OF INVENTION: SUBUNIT ANALOGS

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/435,605A

FILING DATE: 05-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657

REFERENCE/DOCKET NUMBER: A-196B

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-435-605A-9

alignment_scores:

Quality: 12.00 Length: 12

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-435-605A-9 ..

Align seg 1/1 to: US-08-435-605A-9 from: 1 to: 138

213 GltutyrGlnserIyasnSerArgGlnIlepheSer 224
|||||

55 GAATACCAATCTAAAGTTAAAGACAAATATATTTCA 90

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-23

seq_documentation_block:

Sequence 23, Application US/08449045C

Patent No. 5770203

GENERAL INFORMATION:

APPLICANT: Burnette, Neal W.

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

TITLE OF INVENTION: SUBUNIT ANALOGS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; 08-449-045C-23

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-23  ..

Align seg 1/1 to: US-08-449-045C-23 from: 1 to: 50

179 TrrlEhishisAlaProgingIyCysgIyAsn 189
|||||
3 TGGATTCATCATGCACCGCAGGGTGTGGGAAT 35

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-24

seq_documentation_block:
; Sequence 24, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; 08-449-045C-24

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-24  ..

Align seg 1/1 to: US-08-449-045C-24 from: 1 to: 50

179 TrrlEhishisAlaProgingIyCysgIyAsn 189
|||||
3 TGGATTCATCATGCACCGCAGGGTGTGGGAAT 35

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-35

seq_documentation_block:
; Sequence 35, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
; 08-435-605A-35

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
```



```
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,120
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,003
FILING DATE: 11-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
TELEPHONE: (510) 601-2708
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-823-120-36
```

```
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-823-120-36 ..
Align seg 1/1 to: US-08-823-120-36 from: 1 to: 28
```

```
8 Alaaspsarargproaspglu 15
|||||
2 GCAGATCTAGACCTCTGATGAA 25
```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-403-584-5

seq_documentation_block:

```
; Sequence 5, Application US/08403584
; Patent No. 5631010
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; TITLE OF INVENTION: DELETION MUTANTS AS VACCINES
; TITLE OF INVENTION: FOR CHOLERA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,584
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,388
; FILING DATE:
; APPLICATION NUMBER: US/07/909,382
```

```
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00742/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 32
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-403-584-5
```

```
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-403-584-5 ..
Align seg 1/1 to: US-08-403-584-5 from: 1 to: 32
```

```
217 LysValLysArgGlnIlePheSer 224
```

```
|||||
6 AAAGTTAAAGACAAATATTTTCA 29
```

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-367-115-5

seq_documentation_block:

```
; Sequence 5, Application US/08367115
; Patent No. 5874088
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; APPLICANT: Beattie, David
; APPLICANT: Killen, Kevin
; APPLICANT: Lu, Yichen
; TITLE OF INVENTION: DELETION MUTANTS AS VACCINES FOR CHOLERA
; FILE REFERENCE: 00742/002002
; CURRENT APPLICATION NUMBER: US/08/367,115
; CURRENT FILING DATE: 1995-01-05
; EARLIER APPLICATION NUMBER: PCT/US93/06270
; EARLIER FILING DATE: 1993-07-01
; EARLIER APPLICATION NUMBER: 07/909,382
; EARLIER FILING DATE: 1992-07-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-08-367-115-5
```

```
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-367-115-5 ..
Align seg 1/1 to: US-08-367-115-5 from: 1 to: 32
```

```
217 LysValLysArgGlnIlePheSer 224
```

```
|||||
6 AAAGTTAAAGACAAATATTTTCA 29
```



```
seq_name: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:PCT-US93-06270-5
;
; seq_documentation_block:
;   Sequence 5, Application PC/TUS9306270
;   GENERAL INFORMATION:
;     APPLICANT: Mekalnos, John J.
;     TITLE OF INVENTION: DELETION MUTANTS AS VACCINES FOR
;     TITLE OF INVENTION: CHOLERA
;     NUMBER OF SEQUENCES: 6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Fish & Richardson
;       STREET: 225 Franklin Street
;       CITY: Boston
;       STATE: Massachusetts
;       COUNTRY: U.S.A.
;       ZIP: 02110-2804
;
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5" Diskette, 1.44 MB
;     COMPUTER: IBM PS/2 Model 50z or 55sx
;     OPERATING SYSTEM: MS-DOS (Version 5.0)
;     SOFTWARE: Wordperfect (Version 5.1)
;     CURRENT APPLICATION NUMBER: PCT/US93/06270
;     FILING DATE: 19930701
;     CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Freeman, John W.
;       REGISTRATION NUMBER: 29,066
;       REFERENCE/DOCKET NUMBER: 00742/007001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (617) 542-5070
;       TELEFAX: (617) 542-8906
;       TELEX: 200154
;     INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 32
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;       PCT-US93-06270-5
;
; alignment_scores:
;   Quality: 8.00      Length: 8
;   Ratio: 1.000      Gaps: 0
;   Percent Similarity: 100.000      Percent Identity: 100.000
;
; alignment_block:
;   US-09-528-682-1 x PCT-US93-06270-5 ..
;
; Align seq 1/1 to: PCT-US93-06270-5 from: 1 to: 32
;
; 217 LysValLysArgSlnIlePheSer 224
;   |||||||||||||||||||
;   6 AAGCTTAAGACAAATATTTCATCA 29
;
; seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-296-848A-4
;
; seq_documentation_block:
;   Sequence 4, Application US/08296848A
;   Patent No. 6019982
;   GENERAL INFORMATION:
;     APPLICANT: Clements, John D.
;     APPLICANT: Dickinson, Bonny L.
;     TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A
;     TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT
;     NUMBER OF SEQUENCES: 5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: PENNIE & EDMONDS LLP
;       STREET: 1155 Avenue of the Americas
```

```
;
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 10036
;
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/296,848A
;       FILING DATE: 26-AUG-1994
;       CLASSIFICATION: 424
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mistrock, S. Leslie
;       REGISTRATION NUMBER: 18,872
;       REFERENCE/DOCKET NUMBER: 5113-046
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) 790-9090
;       TELEFAX: (212) 869-8864/9741
;       TELEX: 66141 PENNIE
;     INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 45 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: double
;         TOPOLOGY: unknown
;       MOLECULE TYPE: DNA
;     US-08-296-848A-4
;
; alignment_scores:
;   Quality: 8.00      Length: 8
;   Ratio: 1.000      Gaps: 0
;   Percent Similarity: 100.000      Percent Identity: 100.000
;
; alignment_block:
;   US-09-528-682-1 x US-08-296-848A-4 ..
;
; Align seq 1/1 to: US-08-296-848A-4 from: 1 to: 45
;
; 193 ThrIleThrGlyAspThrCysAsn 200
;   |||||||||||||||||||
;   22 ACAATTACAGGTGATCTGTAT 45
;
; seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-360-186-1
;
; seq_documentation_block:
;   Sequence 1, Application US/09360186
;   Patent No. 6262343
;   GENERAL INFORMATION:
;     APPLICANT: Staskawicz, et al.
;     TITLE OF INVENTION: Bst2 Resistance Gene
;     FILE REFERENCE: 50687
;     CURRENT APPLICATION NUMBER: US/09/360,186
;     CURRENT FILING DATE: 1999-07-23
;     EARLIER APPLICATION NUMBER: 60/093,957
;     EARLIER FILING DATE: 1998-07-23
;     NUMBER OF SEQ ID NOS: 9
;     SOFTWARE: Patentin Ver. 2.0
;     SEQ ID NO 1
;     LENGTH: 31491
;     TYPE: DNA
;     ORGANISM: Capsicum annuum
;     US-09-360-186-1
;
; alignment_scores:
;   Quality: 8.00      Length: 8
;   Ratio: 1.000      Gaps: 0
;   Percent Similarity: 100.000      Percent Identity: 100.000
;
; alignment_block:
```


US-09-528-682-1 x US-09-360-186-1/rev ..
Align seg 1/1 to reverse of: US-09-360-186-1 from: 1 to: 31491

114 SerialaLeuGlyIleProTyr 121
|||||
6271 TCTGGCTAGGTGATACCCAC 6248

seq_name: /cgn2_6/ptodata/2/1na/5A.COMB.seq:US-07-932-379A-11

seq_documentation_block:
; Sequence 11, Application US/07932379A
; Patent No. 5468852
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting
; TITLE OF INVENTION: Bacteria and Detection Method
; TITLE OF INVENTION: Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,379A
; FILING DATE: 19920819
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 1327-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..22
; OTHER INFORMATION: /label=oligonucleotide
; ;
; OTHER INFORMATION: /note="Identification method S"
; ;
US-07-932-379A-11

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-07-932-379A-11 ..
Align seg 1/1 to: US-07-932-379A-11 from: 1 to: 22

49 GlnThrGlyPheValArgTyr 55
|||||
2 CAAACCGCGCTTTGTCAGATAT 22

seq_name: /cgn2_6/ptodata/2/1na/5A.COMB.seq:US-07-932-379A-12

seq_documentation_block:
; Sequence 12, Application US/07932379A
; Patent No. 5468852
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting
; TITLE OF INVENTION: Bacteria and Detection Method
; TITLE OF INVENTION: Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,379A
; FILING DATE: 19920819
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 1327-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..22
; OTHER INFORMATION: /label=oligonucleotide
; ;
; OTHER INFORMATION: /note="Identification method S"
; ;
US-07-932-379A-12

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-07-932-379A-12/rev ..

Align seg 1/1 to reverse of: US-07-932-379A-12 from: 1 to: 22

228 SergiuValaspiIetyrasn 234

21 TCAGAGCTTGACATATATATAC 1

seq_name: /cgn2_6/ptodata/2/lna/5a_COMB.seq:US-08-379-295-11

seq_documentation_block:

; Sequence 11, Application US/08379295

; Patent No. 5516898

; GENERAL INFORMATION:

; APPLICANT: Ohashi, Tetsuo

; APPLICANT: Toda, Jun

; APPLICANT: Fukushima, Shigeru

; APPLICANT: Ozaki, Hiroko

; APPLICANT: Nishimura, Nasyuki

; APPLICANT: Yamagata, Koichi

; TITLE OF INVENTION: Oligonucleotides for Detecting

; TITLE OF INVENTION: Bacteria and Detection Method

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 301 N. Washington St.

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22046-3487

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/379,295

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/932,379A

; FILING DATE: 19-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 1327-106P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-241-1300

; TELEFAX: 703-241-2848

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli H10407

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..22

; OTHER INFORMATION: /label= oligonucleotide

; OTHER INFORMATION: /note= "identification method s"

; US-08-379-295-11

alignment_scores:

Quality: 7.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-379-295-11 ..

Align seg 1/1 to: US-08-379-295-11 from: 1 to: 22

49 GlnthrglyphevalArgTyr 55

2 CAACCGCGCTTGTCAGATAT 22

seq_name: /cgn2_6/ptodata/2/lna/5a_COMB.seq:US-08-379-295-12

seq_documentation_block:

; Sequence 12, Application US/08379295

; Patent No. 5516898

; GENERAL INFORMATION:

; APPLICANT: Ohashi, Tetsuo

; APPLICANT: Toda, Jun

; APPLICANT: Fukushima, Shigeru

; APPLICANT: Ozaki, Hiroko

; APPLICANT: Nishimura, Nasyuki

; APPLICANT: Yamagata, Koichi

; TITLE OF INVENTION: Oligonucleotides for Detecting

; TITLE OF INVENTION: Bacteria and Detection Method

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 301 N. Washington St.

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22046-3487

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/379,295

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/932,379A

; FILING DATE: 19-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 1327-106P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-241-1300

; TELEFAX: 703-241-2848

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli H10407

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..22

; OTHER INFORMATION: /label= oligonucleotide


```
; OTHER INFORMATION: /note="identification method s"
; US-08-379-295-12

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-379-295-12/rev  ..

Align seg 1/1 to reverse of: US-08-379-295-12 from: 1 to: 22

228 SerGIuValasPIeTYrAsn 234
|||||
21 TCAGAGGTTGACATATATAC 1

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-296-11

seq_documentation_block:
; Sequence 11, Application US/08379296
; tent No. 5525718
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting Bacteria
; TITLE OF INVENTION: and Detection Method Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,296
; FILING DATE: 27-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,379
; FILING DATE: 19-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 2036-102P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY:
```

```
; LOCATION: 1..22
; OTHER INFORMATION: /label=oligonucleotide
; OTHER INFORMATION: /note="identification method s"
; US-08-379-296-11

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-379-296-11  ..

Align seg 1/1 to: US-08-379-296-11 from: 1 to: 22

49 GlnThrGlyPheValArgTyr 55
|||||
2 CAACCGCGCTTGTGCAGATAT 22

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-296-12

seq_documentation_block:
; Sequence 12, Application US/08379296
; Patent No. 5525718
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting Bacteria
; TITLE OF INVENTION: and Detection Method Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,296
; FILING DATE: 27-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,379
; FILING DATE: 19-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 2036-102P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
```


FEATURE:
NAME/KEY: -
LOCATION: 1..22
OTHER INFORMATION: /label=oligonucleotide
OTHER INFORMATION: /note="identification method S"
US-08-379-296-12

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-379-296-12/rev ..

Align seg 1/1 to reverse of: US-08-379-296-12 from: 1 to: 22

228 SerGluValAspIleTyrAsn 234

|||||
21 TCAGAGGTGCATATATAC 1

seq_name: /cgn2.6/ptodata/2/1na/5A.COMB.seq:US-08-328-710A-26

seq_documentation_block:

Sequence 26, Application US/08328710A

Patent No. 5795717

GENERAL INFORMATION:

APPLICANT: Nakayama, Tomoko

APPLICANT: Tada, Jun

APPLICANT: Fukushima, Shigeru

APPLICANT: Ohashi, Tetsuo

TITLE OF INVENTION: Oligonucleotides for detecting bacteria and detection

TITLE OF INVENTION: process

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch and Birch

STREET: PO Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,710A

FILING DATE: 25-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weiner, Marc S

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 1422-202P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Vibrio cholerae

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-328-710A-26/rev ..

Align seg 1/1 to reverse of: US-08-328-710A-26 from: 1 to: 22

122 SerGluIleTyrGlyTyrTyr 128

|||||
21 TCCCAATATATATGATGATAT 1

seq_name: /cgn2.6/ptodata/2/1na/6B.COMB.seq:US-09-311-260-134

seq_documentation_block:

Sequence 134, Application US/09311260

Patent No. 6214555

GENERAL INFORMATION:

APPLICANT: Leushner, James

APPLICANT: Hul, May

APPLICANT: Dunn, James M.

APPLICANT: Lacroix, Jean-Michel

TITLE OF INVENTION: METHOD, COMPOSITIONS AND KIT FOR DETECTION OF

TITLE OF INVENTION: MICROORGANISMS AND BI-DIRECTIONAL SEQUENCING OF NUCLEIC ACI

TITLE OF INVENTION: POLYMERS

NUMBER OF SEQUENCES: 189

CORRESPONDENCE ADDRESS:

ADDRESSEE: Opdedahl & Larson LLP

STREET: P.O. Box 5270

CITY: Frisco

STATE: CO

COUNTRY: US

ZIP: 80443-5270

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage

OPERATING SYSTEM: IBM compatible

SOFTWARE: Word Perfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/311,260

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Larson, Marina T.

REGISTRATION NUMBER: 32,038

REFERENCE/DOCKET NUMBER: VGEN.P-058-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (970) 668-2050

TELEFAX: (970) 668-2082

TELEX:

INFORMATION FOR SEQ ID NO: 134:

SEQUENCE CHARACTERISTICS:

LENGTH: 22

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

HYPOHETICAL: no

ANTI-SENSE: yes

FRAGMENT TYPE: Internal

US-09-311-260-134

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-311-260-134 ..

Align seg 1/1 to: US-09-311-260-134 from: 1 to: 22

7 ArgalaasSerArgProPro 13
|||||
1 CGGCGAGATTCTAGACCTCT 21

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-968-046-26

seq_documentation_block:
; Sequence 26, Application US/08968046
; Patent No. 6218110
; GENERAL INFORMATION:
; APPLICANT: Nakayama, Tomoko
; APPLICANT: Tada, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ohashi, Tetsuo
; TITLE OF INVENTION: Oligonucleotides for detecting bacteria
; TITLE OF INVENTION: and detection process
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: PO Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,046
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/328,710
; FILING DATE: 25-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-202P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Vibrio cholerae
; US-08-968-046-26

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-968-046-26/rev ..

Align seg 1/1 to reverse of: US-08-968-046-26 from: 1 to: 22

122 SerGlnIleTyrGlyTyr 128
|||||
21 TCCCAATATATGATGCTAT 1

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-363-585-12

seq_documentation_block:
; Sequence 12, Application US/08363585
; Patent No. 5683872
; GENERAL INFORMATION:
; APPLICANT: Rudert, William A.
; APPLICANT: Trucco, Massimo
; TITLE OF INVENTION: Polymers of Oligonucleotide Probes
; TITLE OF INVENTION: As The Bound Ligands For Use In Reverse
; TITLE OF INVENTION: Dot Blots
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: University of Pittsburgh
; STREET: Office of Intellectual Property
; STREET: 911 William Pitt Union
; CITY: Pittsburgh
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 15260
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5-1/4" low density diskette
; COMPUTER: IBM PC or compatibles
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,585
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/786,228
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frederick H. Coleman; Mary-Elizabeth Buckles
; REGISTRATION NUMBER: 28,061; 31,907
; REFERENCE/DOCKET NUMBER: 92-232
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 412/288-4164
; TELEFAX: 412/288-3063
; TELEX: 277871
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: synthetic DNA
; PUBLICATION INFORMATION:
; AUTHORS: Rudert, W.A.
; AUTHORS: Trucco, M.
; TITLE: A No. 5683872el Approach to Rapid HLA Class II
; JOURNAL: Molecular Typing
; JOURNAL: HLA 1991
; VOLUME: 2
; PAGES: 352-356
; DATE: 1992
; RELEVANT RESIDUES IN SEQ ID NO: 12: 1 to 27
; US-08-363-585-12

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-363-585-12 ..

Align seg 1/1 to: US-08-363-585-12 from: 1 to: 27

18 ArgSerGlyGlyLeuMetPro 24
|||||
5 AGATCTGGCGGCTGATCCG 25

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-142-355B-11

seq_documentation_block:

Sequence 11, Application US/09142355B
Patent No. 6184012
GENERAL INFORMATION:
APPLICANT: Neri, Dario
APPLICANT: Demattis, Salvatore
APPLICANT: Huber, Adrain
APPLICANT: Viti, Francesca
APPLICANT: Tawfik, Dan. S.
APPLICANT: Winter, Gregory Paul
TITLE OF INVENTION: Isolation of Enzymes
FILE REFERENCE: 2224/0665
CURRENT APPLICATION NUMBER: US/09/142,355B
CURRENT FILING DATE: 1998-09-04
PRIOR APPLICATION NUMBER: GB 9608540.2
PRIOR FILING DATE: 1996-04-25
PRIOR APPLICATION NUMBER: PCT/GB97/01153
PRIOR FILING DATE: 1997-04-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 42
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer
US-09-142-355B-11

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-142-355B-11/rev ..

Align seg 1/1 to reverse of: US-09-142-355B-11 from: 1 to: 42

61 SerThrSerLeuSerLeuArg 67
|||||
35 AGTACACGCTCAGCTCCG 15

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-276-452A-65

seq_documentation_block:

Sequence 65, Application US/08276452A
Patent No. 5646029
GENERAL INFORMATION:
APPLICANT: Chen, Chao-Guang
APPLICANT: Mau, Shiao-Lim
APPLICANT: Du, He
APPLICANT: Gane, Alison M
APPLICANT: Bacic, Antony
APPLICANT: Clarke, Adrienne E
TITLE OF INVENTION: Plant Arabidnolactan Protein (AGP) Genes
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,452A
FILING DATE: 18-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 27-91A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
TELEX: 49617824

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 85 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: misc_feature

LOCATION: 1..19

OTHER INFORMATION: /note= "T7 promoter sequence"

NAME/KEY: misc_feature

LOCATION: 68..85

OTHER INFORMATION: /note= "Adaptor sequence is equal

OTHER INFORMATION: to nucleotides 444-461 of the NaAGP1 cDNA"

US-08-276-452A-65

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-276-452A-65 ..

Align seg 1/1 to: US-08-276-452A-65 from: 1 to: 85

211 LeuArgGluTyrGlnSerLys 217
|||||

64 TTGCGGGAGTATCAGTCATAA 84

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-798-744-65

seq_documentation_block:

Sequence 65, Application US/08798744
Patent No. 5830747
GENERAL INFORMATION:
APPLICANT: Chen, Chao-Guang
APPLICANT: Mau, Shiao-Lim
APPLICANT: Du, He
APPLICANT: Gane, Alison M
APPLICANT: Bacic, Antony
APPLICANT: Clarke, Adrienne E
TITLE OF INVENTION: Plant Arabidnolactan Protein (AGP) Genes
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,744
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/276,452
; FILING DATE: 18-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 27-91A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; TELEX: 49617824
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..19
; OTHER INFORMATION: /note= "T7 promoter sequence"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 68..85
; OTHER INFORMATION: /note= "Adaptor sequence is equal
; OTHER INFORMATION: to nucleotides 444-461 of the NaAgp1 cDNA"
; US-08-798-744-65

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-798-744-65  ..

Align seg 1/1 to: US-08-798-744-65 from: 1 to: 85

211 LeuArgIuTyrgInserlys 217
|||||
64 TTGCGGAGATCACTCAAAA 84

seq_name: /cgn2_6/plodata/2/ina/5A_COMB.seq:US-09-060-756-266
documentation_block:
; sequence 266, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 266
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
```

```

; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
; US-09-060-756-266

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-060-756-266  ..

Align seg 1/1 to: US-09-060-756-266 from: 1 to: 217

63 SerLeuSerLeuArgSerAla 69
|||||
10 TCCTTATCGCTCGCTCTGCA 30

seq_name: /cgn2_6/plodata/2/ina/5A_COMB.seq:US-08-471-780C-95
documentation_block:
; sequence 95, Application US/08471780C
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,780C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potler, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-471-780C-95

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
```


Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-780C-95/rev ..

Align seg 1/1 to reverse of: US-08-471-780C-95 from: 1 to: 433

62 ThrsSerleuserleuAgsr 68
|||||
340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-467-282B-95

seq_documentation_block:

; Sequence 95, Application US/08467282B
; Patent No. 3800988

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESS: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467, 282B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106, 944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-467-282B-95

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Align seg 1/1 to reverse of: US-08-467-282B-95 from: 1 to: 433

62 ThrsSerleuserleuAgsr 68

|||||
340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-471-282A-95

seq_documentation_block:

; Sequence 95, Application US/08471282A

; Patent No. 5840853

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESS: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471, 282A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106, 944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-471-282A-95

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Align seg 1/1 to reverse of: US-08-471-282A-95 from: 1 to: 433

62 ThrsSerleuserleuAgsr 68
|||||
340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-466-710C-95

seq_documentation_block:

; Sequence 95, Application US/08466710C

; Patent No. 5874541


```
/
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/466,710C
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/106,944
/ FILING DATE: 17-AUG-1993
/ APPLICATION NUMBER: FR 92402326.0
/ FILING DATE: 21-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 93401310.3
/ FILING DATE: 21-MAY-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Potter, Jane E.R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 04958.0008-00000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/ US-08-466-710C-95

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

ment_block:
09-528-682-1 x US-08-466-710C-95/rev ..

Align seg 1/1 to reverse of: US-08-466-710C-95 from: 1 to: 433

62 ThrsSerLeuSerLeuArgSer 68
|||||
340 ACCTCCTGTGCGCTAAGTCG 320

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-471-284B-95

seq_documentation_block:
/ Sequence 95, Application US/08471284B
/ Patent No. 6005079
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
```

```
/
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,284B
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/106,944
/ FILING DATE: 17-AUG-1993
/ APPLICATION NUMBER: FR 92402326.0
/ FILING DATE: 21-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 93401310.3
/ FILING DATE: 21-MAY-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Potter, Jane E.R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 04958.0008-00000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/ US-08-471-284B-95

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-471-284B-95/rev ..

Align seg 1/1 to reverse of: US-08-471-284B-95 from: 1 to: 433

62 ThrsSerLeuSerLeuArgSer 68
|||||
340 ACCTCCTGTGCGCTAAGTCG 320

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-468-739C-95

seq_documentation_block:
/ Sequence 95, Application US/08468739C
/ Patent No. 6015695
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
```


SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,739C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-468-739C-95

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-468-739C-95/rev ..

Align seg 1/1 to reverse of: US-08-468-739C-95 from: 1 to: 433

62 ThrsEusSerLeuArgSer 68
|||||
340 ACCTCTTGTGCTAAGTCG 320

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-471-780C-96
seq_documentation_block:
Sequence 96, Application US/08471780C
Patent No. 5759808
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,780C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-471-780C-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-780C-96/rev ..

Align seg 1/1 to reverse of: US-08-471-780C-96 from: 1 to: 449

62 ThrsEusSerLeuArgSer 68
|||||
352 ACCTCTTGTGCTAAGTCG 332

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-467-282B-96

seq_documentation_block:
Sequence 96, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,282B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-467-282A-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-467-282B-96/rev ..

Align seg 1/1 to reverse of: US-08-467-282B-96 from: 1 to: 449

62 ThrsEulerleuArgSer 68
|||||
352 ACGTCCTGTCGCTAAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-471-282A-96

seq_documentation_block:
Sequence 96, Application US/08471282A
Patent No. 5840853
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-471-282A-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-471-282A-96/rev ..

Align seg 1/1 to reverse of: US-08-471-282A-96 from: 1 to: 449

62 ThrsEulerleuArgSer 68
|||||
352 ACGTCCTGTCGCTAAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-466-710C-96

seq_documentation_block:
Sequence 96, Application US/08466710C
Patent No. 5874541
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,710C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-466-710C-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-466-710C-96/rev ..

Align seg 1/1 to reverse of: US-08-466-710C-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||

352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-471-284B-96

seq_documentation_block:

; Sequence 96, Application US/08471284B
; Patent No. 6005079

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,284B
FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potier, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958, 0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO:

96:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-471-284B-96

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-284B-96/rev ..

Align seg 1/1 to reverse of: US-08-471-284B-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||

352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-468-739C-96

seq_documentation_block:

; Sequence 96, Application US/08468739C
; Patent No. 6015695

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,739C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potier, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958, 0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO:

96:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-468-739C-96

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-468-739C-96/rev ..

Align seg 1/1 to reverse of: US-08-468-739C-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||

352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-050-259B-17

seq_documentation_block:

; Sequence 17, Application US/08050259B


```
; Patent No. 6127169
; GENERAL INFORMATION:
; APPLICANT: Livio Mallucci
; APPLICANT: Valerie Wells
; TITLE OF INVENTION: CELL GROWTH INHIBITORS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; City: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,259B
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB91/01898
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9023907.0
; FILING DATE: 02-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreyer, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58324/WMD/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-050-259B-17

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

nment_block:
09-528-682-1 x US-08-050-259B-17/rev ..

Align seg 1/1 to reverse of: US-08-050-259B-17 from: 1 to: 497

161 GlyTyrArgLeuAlaGlyPhe 167
|||||
446 GGCTACAGCGCTGGCTTC 426

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-770-544-19

seq_documentation_block:
; Sequence 19, Application US/08770544
; Patent No. 5907085
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Ling, Kai-Shu
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; TITLE OF INVENTION: THEIR USES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
```

```
; STREET: Clinton Square, P.O. Box 1051
; City: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,544
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60009008
; FILING DATE: 21-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-770-544-19

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-770-544-19/rev ..

Align seg 1/1 to reverse of: US-08-770-544-19 from: 1 to: 534

59 TyrValSerThrSerLeuSer 65
|||||
501 TACCTTCACACTCTGTCT 481

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-133-804-3

seq_documentation_block:
; Sequence 3, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..758
; OTHER INFORMATION: /product= "26-10 sfv" with
; C-terminal Gly4-Cys"
;
; US-08-133-804-3

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-133-804-3  ..

Align seg 1/1 to: US-08-133-804-3 from: 1 to: 779

15 GlutIeLySArGserGlyGly 21
|||||
729 GAGATTAAACGTCGCGGGCA 749

seq_name: /cgn2.6/pdata/2/1na/5A.COMB.seq:US-08-461-838-3

seq_documentation_block:
; Sequence 3, Application US/08461838
; Patent No. 5753204
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Rind, David B.
; TITLE OF INVENTION: Blosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,838
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..758
; OTHER INFORMATION: /product= "26-10 sfv" with
; C-terminal Gly4-Cys"
;
; US-08-461-838-3

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-461-838-3  ..

Align seg 1/1 to: US-08-461-838-3 from: 1 to: 779

15 GlutIeLySArGserGlyGly 21
|||||
729 GAGATTAAACGTCGCGGGCA 749

```


us-09-528-682-1.01ip2n.rml

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,045C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 07/694,733
APPLICATION NUMBER: 02-MAY-1991
FILING DATE: 02-MAY-1991
APPLICATION NUMBER: US 08/271,222
FILING DATE: 06-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196C
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..777
US-08-449-045C-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-1 ..
US-09-528-682-1 x US-08-449-045C-1

Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGACGAGTACTCAATGATATTCACACCTTATGATCATGC 188
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGAGAACCTCAGACGGGATTGTGTTAGG 216

seq_name: /cgn2_6/prodata/2/lna/backfiles1.seq:5223610-1

seq_documentation_block:
; Patent No. 5223610
; APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
; TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
; HORMONE PROMOTER
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,852
; FILING DATE: 18-MAY-1990
; SEQ ID NO: 1:
; LENGTH: 2020
; 5223610-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x 5223610-1 ..

```


gb_pat.:A06082	163.00	274.56	6.3e-07	791	A06082	PERK620	HindII-fragments
gb_pat.:A06083	163.00	274.56	6.3e-07	791	A06083	PERK620	HindII-fragments
gb_da:VIBT0X	157.00	268.15	1.4e-06	512	K01170	Vibrio cholerae	tox A
gb_da:VIBCTRA2	150.00	262.63	2.9e-06	251	J01847	Vibrio cholerae	cholerae
gb_da:ECOPH1E65	117.00	192.45	0.0234	901	M61015	Escherichia coli	heatt
gb_pat.:B069697	109.50	168.72	0.4919	2724	A69697	Sequence 15	from Pat
gb_pat.:BD009884	109.50	168.72	0.4919	2724	BD009884	Recombinant toxin	
gb_da:VIBCTA1	108.00	168.62	0.0383	255	J01846	Vibrio cholerae	cholerae
gb_pat.:AX020906	99.50	157.85	1.98	1344	AX020906	Sequence 2	from Pat
gb_in:AI114440	97.50	157.03	2.19	1461	AI114440	Caenorhabditis elegans	
gb_da:SC6A5	96.00	140.20	3.66	4501	AI049485	Streptomyces coelic	
gb_da:BEF00XP8B	96.00	131.82	55.86	10854	AI0001886	Deinococcus radiodur	
gb_da:AE009105	96.00	130.01	70.45	13147	AE009105	Agrobacterium tum	
gb_da:AE008070	96.00	129.99	70.59	13147	AE008070	Agrobacterium tum	
gb_pl:AE003314	96.00	114.41	520.84	67580	AE003314	Oryza sativa	genom
gb_da:MMV016	94.50	113.97	551.27	53662	AI019441	Mycobacterium tub	c
gb_in:CNS0706A	94.50	101.71	2.7e+03	194349	AL1590443	chromosome III c	
gb_in:AE030305	94.00	137.87	25.72	3973	AB030305	Pteris rapae	mRNA
gb_pat.:E51095	94.00	137.87	25.72	3973	E51095	Apoptosis-inducing p	
gb_da:BE2A5366	93.00	151.24	4.62	810	AJ245366	Bordetella pertussis	810
gb_da:BE2A5367	93.00	151.24	4.62	810	AJ245367	Bordetella pertussis	810
gb_da:BEA006157	93.00	147.57	7.41	1192	AJ006157	Bordetella pertussis	
gb_da:BEA006153	93.00	147.47	7.50	1204	AJ006153	Bordetella pertussis	
gb_da:BEA006157	93.00	147.02	7.95	1262	AJ006157	Bordetella pertussis	
gb_da:BE27363	93.00	147.02	7.95	1263	AJ007363	Bordetella pertussis	
gb_pat.:BE7364	93.00	147.02	7.95	1263	AJ007364	Bordetella pertussis	
gb_pat.:E03380	93.00	143.03	13.25	1919	I03380	Sequence 5	from Patete
gb_in:AE037676	93.00	136.74	29.70	3716	AB037676	Pteris brassicae m	
gb_pat.:E51096	93.00	136.74	29.70	3716	E51096	Apoptosis-inducing p	
gb_da:BEH032C8	92.50	101.95	2.6e+03	131060	AL132671	Human DNA sequenc	
gb_da:AL646079	92.50	97.33	4.7e+03	2436	AL646079	Ralstonia solanum	
gb_ro:RAEPT93	91.50	128.13	24.87	11596	L19658	Rattus norvegicus	FR
gb_da:AE007993	90.50	120.29	209.37	11596	AE007993	Agrobacterium tum	
gb_da:AE009027	90.50	120.29	245.18	13197	AE009027	Agrobacterium tum	
gb_pr:HS4983	90.50	100.49	3.1e+03	105570	AL160451	Homo sapiens chr1	
gb_hlg:AL550046_1	90.50	100.10	3.3e+03	110000	Conjunction (2 of 4) of		
gb_hlg:AL550046_2	90.50	100.10	3.3e+03	110000	Conjunction (3 of 4) of		
gb_pr:AL590076	90.50	99.33	3.6e+03	119330	AL590076	Human DNA sequenc	
gb_pr:AP001347	90.50	98.16	4.2e+03	134926	AP001347	Homo sapiens gen	
gb_da:AL50318	90.50	94.23	6.9e+03	203833	AL450318	Homo sapiens chr1	
gb_pr:AP001660	90.50	89.36	1.3e+04	340000	AP001660	Homo sapiens gen	
gb_da:AF177946	89.50	130.07	69.91	3925	AF177946	Streptomyces coeli	
gb_da:SC6E6	89.50	111.19	78.97	26	AL359989	Streptomyces coeli	
gb_da:BE2A5368	89.00	144.21	11.40	810	AJ245368	Bordetella pertussis	810
gb_da:BEPTOX51	89.00	139.59	20.61	1316	X16347	Bordetella pertussis	
gb_pat.:A13359	89.00	128.53	85.20	4208	A13359	B.pertussis 10556	TO
gb_da:AE006934	89.00	116.80	380.33	14430	AE006934	Mycobacterium tub	
gb_da:MMV0034	89.00	112.44					

gb-pr:BC000650	88.00	135.73	33.82	1641	BC000650 Homo sapiens, Simla	gb-pat:144694	83.50	125.65	123.23	2061	I44694 Sequence 5 from pate
gb-pr:AK021559	88.00	135.12	36.58	1750	AK021559 Homo sapiens cDNA FL	gb-in:AY051462	83.50	122.86	176.32	2764	AY051462 Drosophila melanog
gb-pr:HS0802252	88.00	128.78	82.49	3407	AL137520 Homo sapiens mRNA: c	gb-ba:BC000650	83.50	119.39	274.83	3976	AY09024 Bacillus cereus merr
gb-ba:BP000650	88.00	125.25	129.71	4336	ML6492 B.bronchiseptica petus	gb-ba:BC000650	83.50	119.03	288.14	4133	M60446 Bacillus sphaericus
gb-pat:AR166753	88.00	118.80	226.52	9717	ARI66753 Sequence 1 from paten	gb-pat:AR166753	83.50	119.03	288.14	4133	A44498 Sequence 1 from pate
gb-ba:AE008681	88.00	118.37	313.30	10165	AE008681 Rickettsia coccidi	gb-pat:AE008681	83.50	119.03	288.14	4133	E03504 DNA encoding insecti
gb-hg:LMF1C8R31.18	88.00	96.10	5.5e+03	10556	AL512786 Homo sapiens chrom	gb-ba:AE009510	83.50	117.36	356.90	4925	AF397025 Brucella abortus
gb-hg:AC009782	88.00	95.71	5.7e+03	110000	Continuation (19 of 22) of I	gb-ba:AE009510	83.50	117.36	356.90	4925	AF397025 Brucella abortus
gb-hg:AC0011772	88.00	94.35	6.8e+03	126851	AC0092815 Homo sapiens chrom	gb-hg:AC0011772	83.50	83.59	2.7e+04	171122	AL158039 Homo sapiens chr
gb-hg:AC0048365	88.00	92.57	8.6e+03	128951	AC0011772 Homo sapiens chrom	gb-hg:AC0048365	83.50	82.74	3.2e+04	197019	AL159383 Homo sapiens chr
gb-pr:AL596326	88.00	91.28	1.0e+04	178076	AC0048365 Homo sapiens clone	gb-pr:AL596326	83.50	82.24	3.2e+04	253431	AL045064 Plasmid DNA sequen
gb-pr:AC018795	88.00	90.61	1.1e+04	187836	AC018795 Homo sapiens chrom	gb-ba:AC018795	83.50	78.45	5.2e+04	338100	AL045064 Plasmid DNA sequen
gb-ba:AF110332	87.50	132.40	7.97	458	AF110332 Mesorhizobium sp. L09	gb-ba:AF110332	83.00	134.23	40.94	762	M33958 Chinese hamster prion
gb-pr:HS082009	87.50	132.40	51.83	2123	AF032882 Homo sapiens forkhead	gb-pr:HS082009	83.00	125.14	131.59	1983	AF032882 Homo sapiens hydrop
gb-in:LM06255	87.50	90.63	1.1e+04	170908	AL121582 Human DNA sequence	gb-in:LM06255	83.00	123.34	155.61	2394	AF032882 Homo sapiens hydrop
gb-hg:AC008620	87.50	89.96	1.2e+04	177648	AL359716 Leishmania major F	gb-pr:AC008620	83.00	123.34	155.61	2394	AF032882 Homo sapiens hydrop
gb-pr:AC022336	87.50	89.12	1.3e+04	183399	AC008620 Homo sapiens chrom	gb-pr:AC022336	83.00	123.34	155.61	2394	AF032882 Homo sapiens hydrop
gb-pr:AF137379	87.50	89.12	1.3e+04	198392	AC022336 Mus musculus chrom	gb-pr:AF137379	83.00	123.16	159.50	2440	AF139572 Mus musculus carti
gb-pr:AF137379	87.50	89.10	1.3e+04	200799	AF137379 Nephroselmis olivac	gb-pr:AF137379	83.00	107.91	1.2e+03	12119	AY070571 Drosophila melanog
gb-hg:AC0090443	87.50	88.97	1.4e+04	200799	AF137379 Nephroselmis olivac	gb-hg:AC0090443	83.00	105.46	1.2e+03	15670	AE007516 Mycobacterium aceto
gb-ba:AE0242847	87.00	128.59	84.51	28889	AF0242847 Bordetella bronchise	gb-ba:AE0242847	83.00	97.23	4.7e+03	37218	Z77162 Mycobacterium tuber
gb-pr:SP0669	87.00	112.72	66.74	15301	AF0242847 Bordetella bronchise	gb-pr:SP0669	83.00	96.63	5.3e+03	40969	AL154151 Caenorhabditis eleg
gb-pr:AC019013	87.00	93.60	7.5e+03	114036	AL035226 S.pombe chromosome 1	gb-pr:AC019013	83.00	86.60	1.4e+04	92118	AC106813 Rattus norvegicus
gb-hg:AC009255	86.50	131.47	1.2e+04	163217	AC009255 Rattus norvegicus F	gb-hg:AC009255	83.00	87.20	2.2e+04	106729	AC009320 Human Chromosome
gb-ba:NM06255	86.50	131.47	58.44	1547	X74217 Z.mays mRNA for beta-D	gb-ba:NM06255	83.00	84.12	2.5e+04	145825	AC005729 Homo sapiens, **
gb-ba:AE002482	86.50	120.98	224.18	5857	X82637 N.meningitidis pdh gene	gb-ba:AE002482	83.00	84.12	2.5e+04	145825	AC005729 Homo sapiens, **
gb-ba:AE002482	86.50	114.18	536.35	11968	AC002482 Neisseria meningitidis	gb-ba:AE002482	83.00	83.18	2.9e+04	162813	AC021833 Homo sapiens chr
gb-ba:AE002482	86.50	85.64	2.1e+04	240050	AL627267 Salmonella enterica	gb-ba:AE002482	83.00	83.09	2.9e+04	162813	AC021833 Homo sapiens chr
gb-pat:AX044032	86.50	128.46	83.95	2435	AF007768 Choriostoma fumifera	gb-pat:AX044032	83.00	82.81	3.0e+04	167819	AC103528 Rattus norvegicus
gb-in:AE007768	86.00	125.81	19.51	3190	AF007768 Choriostoma fumifera	gb-in:AE007768	83.00	82.81	3.0e+04	167819	AC103528 Rattus norvegicus
gb-ba:AE023523	86.00	125.56	144.56	3300	AF023523 Pseudomonas aeruginosa	gb-ba:AE023523	83.00	82.77	3.0e+04	169308	AL1591000 Homo sapiens chr
gb-ba:AE009044	86.00	114.25	539.04	10790	AE009044 Agrobacterium tumefaci	gb-ba:AE009044	83.00	81.33	3.6e+04	178073	AL1590461 Streptomyces coe
gb-ba:AE009044	86.00	114.25	539.04	10790	AE009044 Agrobacterium tumefaci	gb-ba:AE009044	83.00	81.33	3.6e+04	178073	AL1590461 Streptomyces coe
gb-ba:AE004646	86.00	113.90	556.34	11244	AE004646 Pseudomonas aeruginosa	gb-ba:AE004646	83.00	80.81	3.9e+04	208642	AL117339 Human DNA sequen
gb-ba:AE004646	86.00	83.71	2.7e+04	268200	AP000661 Aeterypium pernix g	gb-ba:AE004646	83.00	80.81	3.9e+04	208642	AL117339 Human DNA sequen
gb-com:AY029774	85.50	138.53	78.67	2065	AY029774 Ovis aries prion prote	gb-com:AY029774	83.00	126.97	104.08	1492	M28330 A.oxydatis 6-hydroxy-
gb-com:ASHEMOLYS	85.50	129.15	82.13	2139	X65046 A.sobria DNA sequence 4	gb-com:ASHEMOLYS	83.00	119.61	250.64	3305	L08502 Drosophila melanog
gb-hg:AC009491	85.50	92.79	8.3e+03	194186	AC009491 Rattus norvegicus cl	gb-hg:AC009491	83.00	119.61	250.64	3305	L08502 Drosophila melanog
gb-pr:AC005456	85.50	88.54	1.4e+04	147185	AP004569 Oryza sativa chrom	gb-pr:AC005456	83.00	119.42	273.99	3397	AY074759 Drosophila melanog
gb-hg:AC004772	85.50	88.14	1.5e+04	153506	AP004772 Oryza sativa chrom	gb-hg:AC004772	83.00	119.42	273.99	3397	AY074759 Drosophila melanog
gb-ba:AE027272	85.50	83.43	2.7e+04	245050	AL627272 Salmonella enterica	gb-ba:AE027272	83.00	107.89	1.2e+03	11069	AE004668 Pseudomonas aerug
gb-ba:AE000062	85.50	83.43	2.7e+04	251700	AP000062 Aeterypium pernix g	gb-ba:AE000062	83.00	93.34	7.8e+03	36774	AC019572 Drosophila melan
gb-pat:AX121338	85.00	130.63	65.04	1511	AX121338 Sequence 1254 from Pa	gb-pat:AX121338	82.50	85.04	2.3e+04	122056	P090902 Synecocystis sp.
gb-ov:AE004885	85.00	130.09	69.75	1706	AF128510 Trapaetulus periscus GNH	gb-ov:AE004885	82.50	82.18	3.3e+04	164974	AC011249 Drosophila melan
gb-ba:AE004885	85.00	110.59	84.95	13223	AE004885 Pseudomonas aerugin	gb-ba:AE004885	82.50	82.18	3.3e+04	164974	AC011249 Drosophila melan
gb-ba:SM0591787	85.00	87.49	1.6e+04	149870	AC024579 Homo sapiens chrom	gb-ba:SM0591787	82.50	81.65	3.5e+04	174376	AC092230 Drosophila melan
gb-pat:AX127146	85.00	79.41	4.6e+04	349980	AX127146 Sequence 7062 from	gb-pat:AX127146	82.50	79.49	4.6e+04	181874	AC018064 Drosophila melan
gb-ba:NM06664	84.50	124.57	141.46	2276	X77920 N.meningitidis outer me	gb-ba:NM06664	82.50	77.87	5.8e+04	257662	AE003577 Drosophila melan
gb-ba:AE008779	84.50	103.58	2.1e+03	25184	AE008779 Salmonella typhimur	gb-ba:AE008779	82.50	77.09	6.2e+04	281457	AL108494 Homo sapiens chr
gb-pl:NC158	84.50	95.43	6.3e+03	61843	AL158382 Neutropora crassa DN	gb-pl:NC158	82.50	76.35	6.8e+04	304290	AF242827 Pyrococcus abyssi
gb-pr:AC0589734	84.50	94.43	6.8e+03	65882	AL1589734 Human DNA sequence 4	gb-pr:AC0589734	82.50	76.23	6.9e+04	307990	AC098216 Rattus norvegicus
gb-hg:AL592186	84.50	85.52	2.1e+04	168018	AL592186 Homo sapiens chrom	gb-hg:AL592186	82.50	75.02	8.1e+04	349980	AX041921 Sequence 816 fro
gb-hg:AC013703	84.50	85.04	2.3e+04	176704	AC013703 Homo sapiens clone	gb-hg:AC013703	82.00	134.08	41.77	644	AF113310 Bos javanicus prion
gb-ba:AL646677	84.50	84.40	2.4e+04	189050	AL646677 Ralstonia solanacea	gb-ba:AL646677	82.00	134.02	42.09	648	AF113327 Bos taurus prion pr
gb-com:OAVJ7271	84.00	135.66	34.15	790	AF000734 Ovis aries PVP gene, a	gb-com:OAVJ7271	82.00	133.21	52.04	771	AF113942 Girafla camelopard
gb-vi:AF168005	84.00	120.80	229.49	3762	AF168005 Panna vitus segment 1	gb-vi:AF168005	82.00	132.37	52.04	771	AF113942 Girafla camelopard
gb-ba:AE004815	84.00	111.26	780.50	10254	AF004815 Pseudomonas aerugin	gb-ba:AE004815	82.00	132.37	52.29	774	L07623 Sus scrofa prion prot
gb-ba:AE003391	84.00	105.92	1.5e+03	17971	AB033391 Bacillus circulans B	gb-ba:AE003391	82.00	129.11	79.06	1086	M1585 Rat 6-phosphotriecto
gb-hg:AC012890	84.00	90.01	1.2e+04	95614	AC012890 Drosophila melanog	gb-hg:AC012890	82.00	124.92	135.35	1887	X13579 Rat mRNA for 6-phosp
gb-in:AC009197	84.00	84.52	2.4e+04	170057	AC009197 Drosophila melanog	gb-in:AC009197	82.00	124.92	135.35	1887	X13579 Rat mRNA for 6-phosp
gb-in:AC008182	84.00	83.25	2.8e+04	192350	AC008182 Drosophila melanog	gb-in:AC008182	82.00	123.80	156.29	1898	J04197 Rattus norvegicus 6-
gb-in:AE003623	84.00	80.21	4.2e+04	267547	AE003623 Drosophila melanog	gb-in:AE003623	82.00	120.94	225.31	2561	AY058170 Arabidopsis thalia
gb-pat:ED08061	83.50	125.65	123.23	2061	E08061 DNA encoding variant cy	gb-pat:ED08061	82.00	120.71	232.02	2624	AY058170 Arabidopsis thalia
gb-pat:123834	83.50	125.65	123.23	2061	I23834 Sequence 9 from patent	gb-pat:123834	82.00	120.71	232.02	2624	AY058170 Arabidopsis thalia
gb-pat:143345	83.50	125.65	123.23	2061	I43345 Sequence 5 from patent	gb-pat:143345	82.00	120.54	238.26	2681	X15580 R.norvegicus mRNA fo
gb-ba:BP000650	83.50	125.65	123.23	2061	X64876 B.pertussis timbrial	gb-ba:BP000650	82.00	120.54	238.26	2681	X64876 B.pertussis timbrial

gb_hlg:LMFCHR36_25	79.50	80.76	3.9e+04	110000	Continuation (26 of 36) of 1	gb_da:AE005504	78.50	100.00	3.3e+03	12115	AE005504 Escherichia coli
gb_pr:AC005383	79.50	79.69	4.4e+04	121110	AC005383 Homo sapiens chrom	gb_da:AE007423	78.50	99.32	3.6e+03	13002	AE007423 Streptococcus pneu
gb_hlg:AE004538	79.50	77.88	5.6e+04	148885	AP004538 Oryza sativa chrom	gb_da:AE003997	78.50	98.82	3.8e+03	13714	AE003997 Xylella fastidiosa
gb_hlg:AE004665	79.50	77.29	6.0e+04	158298	AP004665 Oryza sativa chrom	gb_hlg:AC013238	78.50	94.33	6.8e+03	21972	AC013238 Drosophila melan
gb_hlg:AE001646	79.50	75.52	7.6e+04	190777	AC001646 Homo sapiens chrom	gb_da:AE004669	78.50	94.13	7.0e+03	22444	AE004669 Pseudomonas aerug
gb_hlg:AC013746	79.50	75.17	7.9e+04	197872	AC013746 Homo sapiens chrom	gb_pr:AL590293	78.50	83.59	2.7e+04	67911	AL590293 Human DNA sequenc
gb_pr:AC017093	79.50	74.92	8.1e+04	203043	AC017093 Homo sapiens BAC cl	gb_pl:NCB101	78.50	82.51	3.1e+04	76072	AL355927 Neurospora crassa
gb_da:BSB00016	79.50	74.35	8.8e+04	215640	129918 Bacillus subtilis com	gb_hlg:SNED1910	78.50	82.30	3.2e+04	77743	AL449932 Streptococcus pneu
gb_da:AE008220	79.50	74.16	9.0e+04	220060	AF008220 Bacillus subtilis r	gb_hlg:AC091350	78.50	79.64	4.5e+04	102835	AC091350 Rattus norvegicus
gb_da:TAACID4	79.50	71.28	1.3e+05	297750	AL445066 Thermoplasma acidop	gb_hlg:AC021947	78.50	77.51	5.9e+04	128574	AC021947 Homo sapiens cld41
gb_da:AP000030	79.50	70.72	1.4e+05	319000	AP000006 Pyrococcus horikosh	gb_hlg:AC0044793	78.50	76.67	6.9e+04	141579	AC0044793 Homo sapiens chr
gb_pat:AX004006	79.50	69.74	1.6e+05	349980	AX044030 Pyrococcus 109 from F	gb_hlg:AC0079981	78.50	76.20	6.9e+04	146516	AC0079981 Homo sapiens chr
gb_com:AF113940	79.00	129.06	79.54	627	AF113940 Camelus dromedarius pr	gb_pr:AL391117	78.50	75.44	7.6e+04	159940	AL391117 Human DNA sequenc
gb_com:AF113180	79.00	128.57	84.68	660	AF009180 Odocolleus hemionus F	gb_da:AL591873	78.50	73.31	1.0e+05	200050	AL591873 Listeria monococy
gb_com:AF117313	79.00	128.30	87.66	679	AF117313 Gazella subgutturosa F	gb_hlg:AC0984730	78.50	71.99	1.2e+05	229788	AC0984730 Mus musculus chr
gb_pr:CTG075383	79.00	127.83	93.21	714	UT5383 Theropithecus gaudata pri	gb_da:AP000064	78.50	71.27	1.3e+05	247695	AP000064 Aeropyrum pernix14
gb_pr:CTG075385	79.00	127.51	97.05	738	UT5385 Cercopithecus torquatus aty	gb_com:AC002562	78.50	70.44	1.4e+05	270365	AP002562 Escherichia coli
gb_com:AF117315	79.00	127.09	102.37	771	AF009181 Odocolleus hemionus h	gb_in:AE003432	78.50	68.85	1.8e+05	319551	AE003432 Drosophila melan
gb_com:AF1195247	79.00	127.09	102.37	771	AF117315 Capra hircus cell-line	gb_com:AF117309	78.00	127.96	91.57	585	AF117309 Adax nasomaculatus
gb_com:AF195247	79.00	127.09	102.37	771	AF195247 Ovis aries prion prote	gb_com:AF117319	78.00	127.63	93.60	606	AF117319 Capra ibex nubiana
gb_pr:ISS008310	79.00	126.95	104.32	783	U08310 Salimiri schiurens prion F	gb_com:OH097323	78.00	126.62	106.10	60	U97331 Odocolleus hemionus F
gb_com:AF156184	79.00	126.39	112.01	830	AF156184 Odocolleus virginianus	gb_com:AF117323	78.00	126.87	108.06	670	AF117323 Hippoboscus niger F
gb_com:AF156185	79.00	126.39	112.01	830	AF156185 Odocolleus virginianus	gb_com:AF156186	78.00	125.74	121.80	739	AF156186 Odocolleus virginianus
gb_com:OAA00681	79.00	125.01	133.79	960	AF000681 Ovis aries prp gene, a	gb_pr:AF151564	78.00	125.49	125.84	759	U15164 Ateles paniscus x Ateles
gb_da:AF131321	79.00	120.69	232.75	1511	AF311321 Sinorhizobium fredii	gb_com:AB060288	78.00	125.34	128.27	771	AB060288 Ovis aries musimon
gb_ro:AF373861	79.00	119.93	256.66	1637	AF373861 Mus musculus N-methyl	gb_com:AB060289	78.00	125.34	128.27	771	AB060289 Ovis aries musimon
gb_pl:ZEMNRBP	79.00	119.24	280.33	1760	Z26824 Zea mays gene for Ferre	gb_com:AF117316	78.00	125.34	128.27	771	AF117320 Ovis montanus pr
gb_in:IND017042	79.00	110.87	820.67	4232	U17042 Leishmania donovani put	gb_com:AF117320	78.00	125.34	128.27	771	AF117320 Ovis montanus pr
gb_sy:AF072539	79.00	110.68	840.08	4324	AF072539 Cloning vector pCMV7A	gb_com:AF166334	78.00	125.34	128.27	771	AF166334 Ovis canadensis pr
gb_pl:PM055002	79.00	110.68	840.08	4324	U505002 Panicle mosaic virus (F	gb_com:CHFRPR	78.00	125.34	128.27	771	X74758 C.hircus gene for pr
gb_pl:MZEPKPKA2	79.00	108.36	1.4e+03	5519	M58656 Corn pyruvate, orthopos	gb_com:OH025665	78.00	125.34	128.27	771	U25965 Odocolleus hemionus F
gb_vl:AF441333	79.00	107.62	1.2e+03	5964	AF441333 Bean leafroll virus,	gb_com:S82626	78.00	125.34	128.27	771	S82626 Prp-prion protein (at
gb_da:AF041819	79.00	102.66	2.3e+03	11019	AF0041819 Mycobacterium bovis	gb_pat:AX164161	78.00	125.34	128.27	771	AX164161 Sequence 9 from Pat
gb_da:AE000347	79.00	101.56	2.7e+03	11272	AE000347 Escherichia coli K12	gb_pat:AX164163	78.00	125.34	128.27	771	AX164163 Sequence 11 from Pat
gb_da:AE000563	79.00	100.74	3.0e+03	12295	AE000563 Escherichia coli O15	gb_pat:AX164165	78.00	125.34	128.27	771	AX164165 Sequence 13 from Pat
gb_da:AE0007027	79.00	97.37	4.6e+03	17506	AE0007027 Mycobacterium tuberc	gb_com:OAA1735	78.00	125.10	132.14	790	AX100735 Ovis aries prp gene
gb_pr:AL355567	79.00	85.59	2.1e+04	143391	AL355567 Human DNA sequence f	gb_da:AF109343	78.00	123.16	170.84	975	AF109343 Streptomyces griseu
gb_pr:AL355823	79.00	77.29	6.0e+04	143391	AL355823 Homo sapiens clone	gb_da:AF111768	78.00	122.66	180.73	1021	AF117688 Pseudomonas resin
gb_pr:AL357894	79.00	76.62	6.0e+04	158487	AL357894 Human DNA sequence	gb_da:AS025654	78.00	122.31	189.19	1060	AJ252654 Agricultural soil
gb_hlg:AC084796	79.00	76.41	6.7e+04	155824	AC084796 Homo sapiens chrom	gb_pl:YSCATP1A	78.00	116.76	385.49	1899	D63449 Yeast ATRI gene for
gb_hlg:AC0024720	79.00	76.56	6.7e+04	158338	AC0024720 Homo sapiens chrom	gb_pat:EO6817	78.00	116.39	404.16	1974	EO6817 DNA encoding alcohol
gb_pr:AL353586	79.00	76.15	7.0e+04	162695	AL353586 Homo sapiens chrom	gb_pat:IC108050	78.00	116.39	404.16	1974	E08050 CDNA encoding beer y
gb_pr:AC022166	79.00	76.08	7.0e+04	163998	AC022166 Homo sapiens chrom	gb_pat:IC121407	78.00	116.39	404.16	1974	IC121407 Sequence 16 from pat
gb_hlg:AC0009336	79.00	75.15	7.9e+04	180759	AC009336 Pan troglodytes cld	gb_pat:IC15437	78.00	116.39	404.16	1974	IC15437 Sequence 16 from pat
gb_da:AP000063	79.00	74.92	8.2e+04	185300	AP000063 Aeropyrum pernix g	gb_pat:IC173372	78.00	116.39	404.16	1974	IC17372 Sequence 16 from pat
gb_hlg:AC016168	79.00	74.52	8.6e+04	193211	AC016168 Homo sapiens chrom	gb_pat:IC193338	78.00	116.39	404.16	1974	IC19338 Sequence 16 from pat
gb_hlg:AC092508	79.00	74.41	8.7e+04	195411	AC092508 Mus musculus chrom	gb_com:SHPRPA	78.00	114.35	459.80	2194	D38179 Sheep gene for prion
gb_pr:AL51074	79.00	74.04	9.1e+04	203076	AL51074 Human DNA sequence	gb_com:S46825	78.00	114.35	535.97	2446	S46825 prion protein (mank
gb_da:AL667269	79.00	71.91	1.2e+05	254505	AL627269 Salmonella enterica	gb_com:OARPR1	78.00	105.32	962.99	4020	AJ323072 Ovis aries prp gene
gb_da:AP002567	79.00	69.75	1.6e+05	318703	AP002567 Escherichia coli O1	gb_pl:AL062171	78.00	105.63	1.7e+03	6309	AL062171 Neurospora crassa
gb_da:AP002997	79.00	69.43	1.6e+05	328709	AP002997 Mesorhizobium loti	gb_da:DB7026	78.00	101.34	2.7e+03	9390	DB7026 Bacillus stearotherm
gb_da:AP003002	79.00	68.96	1.7e+05	346547	AP003002 Mesorhizobium loti	gb_da:AE000496	78.00	99.29	3.5e+03	11533	AE000496 Agrobacterium tun
gb_ov:AF038161	79.00	134.29	40.67	330	AF038161 Ictalurus punctatus cl	gb_da:AE000124	78.00	99.19	3.7e+03	11929	AE000124 Agrobacterium tun
gb_ov:AF113943	78.50	128.51	85.40	606	AF113943 Lama glama prion prote	gb_da:AE000809	78.00	98.60	4.0e+03	12074	AE000809 Agrobacterium tun
gb_pr:AF076916	78.50	126.33	112.96	762	AF076916 Homo sapiens prion prote	gb_com:OARPR	78.00	90.05	1.2e+04	31412	U67922 Ovis aries prion pr
gb_com:CDPRPACLF	78.50	126.25	114.05	768	Y09760 C.dromedarius prp gene.	gb_hlg:AC012904	78.00	81.48	2.4e+04	56423	AC012904 Drosophila melan
em_hlgO_inlv:AE029300	78.50	123.50	162.42	1026	AC029300 Giardia intestinalis	gb_pl:AB025631	78.00	79.43	4.6e+04	81035	AB025631 Arabidopsis thalli
em_hlgO_inlv:AE029300	78.50	123.13	170.18	1066	AC029300 Giardia intestinalis	gb_pr:AL359510	78.00	79.37	4.6e+04	96474	AL359510 Human DNA sequenc
gb_pat:AX078894	78.50	122.52	184.12	1137	AX078894 Sequence 5 from Paten	gb_hlg:AC0096097	78.00	79.43	4.6e+04	96474	AC0096097 Rattus norvegicus
gb_da:CDHOXINA	78.50	117.39	355.26	1948	X17194 Clostridium difficile c	gb_pr:AL354993	78.00	78.41	5.2e+04	106736	AL354993 Human DNA sequen
gb_pat:AL21365	78.50	117.39	355.26	1948	X17194 Clostridium difficile c	gb_hlg:AC093316_0	78.00	78.12	5.4e+04	110000	AC093316 Mus musculus chr
gb_pat:IL15072	78.50	117.39	355.26	1948	X17194 Clostridium difficile c	gb_hlg:LMFCHR31_12	78.00	78.12	5.4e+04	110000	LMFCHR31_12 Mus musculus chr
gb_da:STR0V8402A	78.50	114.75	498.50	2571	M80215 Streptococcus pneumonia	gb_hlg:LMFCHR31_13	78.00	78.12	5.4e+04	110000	LMFCHR31_13 Mus musculus chr
gb_pat:BD0053822	78.50	108.38	1.1e+03	5020	BD0053822 Polynucleotide of Str	gb_hlg:LMFCHR32_19	78.00	78.12	5.4e+04	110000	LMFCHR32_19 Mus musculus chr
gb_da:ECV10545	78.50	101.87	1.1e+03	5025	Y10545 Cloning vector pGEMO-21	gb_pr:AC005301	78.00	77.81	5.7e+04	113688	AC005301 Homo sapiens chr
gb_da:AE0008484	78.50	101.80	1.2e+03	10029	AE0008484 Streptococcus pneum	gb_pr:AC016515	78.00	77.23	6.1e+04	120831	AC016515 Homo sapiens cld41
gb_da:AE0008930	78.50	100.82	3.0e+03	11107	AE0008930 Methanobacterium the	gb_pr:AL157772	78.00	77.14	6.2e+04	121949	AL157772 Human DNA sequenc

gb-pr:AL136379	-	78.00	76.69	6.5e+04	127816	! AL136379 Human DNA sequence	gb-ba:AP003011	77.50	66.32	2.4e+05	346510	! AP003011 Mesorhizobium 10
gb-pr:AC009655	-	78.00	76.17	7.0e+04	135033	! AC009655 Homo sapiens BAC cl	gb-pat:AX120085	77.50	66.22	2.4e+05	349980	! AX120085 Sequence 18 from Pa
gb-hg:AC009673	+	78.00	75.12	7.4e+04	141455	! AC009673 Trypanosoma brucei	gb-pat:AX120695	77.00	125.34	1.2e+05	641	! AX057695 Sequence 18 from Pa
gb-pr:AC009848	+	78.00	75.51	7.6e+04	144724	! AC009848 Oryza sativa chr10	gb-pat:AB060290	77.00	123.58	160.72	771	! AB060290 Budorcas taxicolor
gb-pr:AC009810	+	78.00	77.86	8.2e+04	156071	! AC009810 Homo sapiens chr10	gb-pat:AX161617	77.00	123.58	160.72	771	! AX161617 Sequence 15 from Pa
gb-hg:AC008407	-	78.00	74.86	8.2e+04	155007	! AC008407 Trypanosoma brucei	gb-pat:AX299244	77.00	123.15	169.16	804	! AX299244 Sequence 1 from Pat
gb-hg:AL513545	-	78.00	74.66	8.4e+04	158297	! AL513545 Homo sapiens chr10	gb-pat:AX299244	77.00	121.49	210.04	960	! AX299244 Ovis aries PEP gene
gb-hg:AC0010565	-	78.00	74.56	8.5e+04	159928	! AC0010565 Drosophila melanog	gb-pat:AX299244	77.00	121.49	210.04	960	! AX299244 Ovis aries PEP gene
gb-hg:AC0024009	-	78.00	73.28	1.0e+05	182662	! AC0024009 Homo sapiens chr10	gb-pat:AX145127	77.00	118.45	310.42	1332	! AR145127 Sequence 27 from p
gb-hg:AC0040991	-	78.00	73.08	1.0e+05	166883	! AC0040991 Homo sapiens chr10	gb-pat:AX133947	77.00	117.95	331.47	1395	! AX133947 Arabidopsis thalia
gb-hg:AC0069091	-	78.00	72.69	1.1e+05	194713	! AC0069091 Homo sapiens chr10	gb-pat:AX133947	77.00	117.85	335.24	1408	! X03146 Klebsiella pneumonia
gb-hg:AC0069098	-	78.00	72.52	1.1e+05	198192	! AC0069098 Mus musculus chr10	gb-pat:AX151114	77.00	115.16	473.10	1867	! Y15114 Anabaena PCC7210 crt
gb-pr:AC0021020	-	78.00	72.30	1.1e+05	202771	! AC0021020 Homo sapiens BAC cl	gb-pr:AC0021020	77.00	114.05	545.19	2097	! B13685 Mouse p10n protein
gb-hg:AC0041335	-	78.00	71.53	1.3e+05	219940	! AC0041335 Homo sapiens chr10	gb-pr:AC0021020	77.00	113.87	558.23	2138	! BC020074 Mus musculus, Sml
gb-hg:AC0041335	-	78.00	71.53	1.3e+05	220107	! AC0041335 Homo sapiens chr10	gb-pr:AC0021020	77.00	113.87	558.23	2138	! BC020074 Mus musculus, Sml
gb-hg:AL606909	+	78.00	71.35	1.3e+05	224061	! AL606909 Mus musculus chr10	gb-ba:AB0451585	77.00	111.58	748.62	2719	! AF051693 Pseudomonas aerugi
gb-hg:AL627425	+	78.00	70.97	1.3e+05	223050	! AL627425 Mus musculus chr10	gb-pr:AC0021020	77.00	110.82	825.61	2946	! AX278949 Cyrtosporogony op
gb-ba:AL627425	+	78.00	70.97	1.3e+05	223050	! AL627425 Mus musculus chr10	gb-pr:AC0021020	77.00	107.88	1.2e+03	4032	! AX123472 Sequence 3388 from
gb-ba:AP000995	+	78.00	69.26	1.7e+05	279160	! AP000995 Thermoplasma volcan	gb-pat:AX137063	77.00	106.17	1.5e+03	4800	! AB021131 Mus musculus RIM2
gb-ba:AP000345	+	78.00	69.16	1.7e+05	281893	! AP000345 Drosophila melanog	gb-pat:AX137063	77.00	105.82	1.6e+03	4980	! AX137063 Sequence 2 from Pa
gb-ba:AP002994	+	78.00	67.17	2.2e+05	347660	! AP002994 Mesorhizobium loti	gb-pr:AC0021020	77.00	103.34	2.2e+03	6466	! D65948 Mus musculus mRNA fo
gb-ba:AB074293	+	77.50	124.99	134.09	729	! AB074293 Photobacterium damela	gb-ba:AY008284	77.00	100.58	3.1e+03	8637	! AY008284 Yersinia enterocol
gb-pat:AX120221	+	77.50	122.73	119.09	924	! AX120221 Sequence 137 from Pat	gb-ba:AY008284	77.00	98.61	3.9e+03	10619	! AE002546 Neisseria meningi
em-hg:AC081030	+	77.50	122.50	184.55	947	! AC081030 Giardia intestinalis c	gb-ba:AE008332	77.00	98.56	4.0e+03	10684	! AE008332 Agrobacterium tum
gb-pat:AC005883	+	77.50	121.79	202.06	1020	! AC005883 Sequence 227 from Pat	gb-ba:AE008332	77.00	91.51	9.8e+03	22388	! AE008332 Agrobacterium tum
em-hg:AC081030	+	77.50	121.47	210.55	1055	! AC081030 Giardia intestinalis c	gb-ba:AE008332	77.00	83.73	2.7e+04	50734	! AC101935 Mus musculus c10
gb-ba:BA000000	+	77.50	113.25	604.22	2502	! BA000000 Giardia intestinalis c	gb-ba:AE008332	77.00	81.73	3.7e+04	56821	! AC101935 Mus musculus c10
gb-ba:BA000000	+	77.50	112.33	680.22	2757	! BA000000 Giardia intestinalis c	gb-ba:AE008332	77.00	79.92	4.3e+04	75703	! AC007722 Homo sapiens chr10
gb-ba:BA000000	+	77.50	112.33	680.22	2757	! BA000000 Giardia intestinalis c	gb-ba:AE008332	77.00	79.92	4.3e+04	75703	! AC007722 Homo sapiens chr10
gb-pat:AL101999	+	77.50	111.84	723.84	2901	! AL101999 Sinorhizobium melioid	gb-ba:AE008332	77.00	76.65	5.5e+04	97137	! AL355112 Homo sapiens chr10
gb-pat:AL101999	+	77.50	107.75	1.2e+03	4460	! AL101999 Sinorhizobium melioid	gb-ba:AE008332	77.00	76.65	5.5e+04	97137	! AL355112 Homo sapiens chr10
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40	1.2e+03	4487	! AE029829 Pseudomonas fluoresc	gb-ba:AE029829	77.00	75.47	7.2e+04	115532	! AL354754 Human DNA sequen
gb-ba:AE029829	+	77.50	107.40									

gb_sts:HSHP076	-	76.50	131.88	55.45	294	AL158939 H. sapiens STS from fld	gb_hcg:AL645966	-	76.50	71.19	1.3e+05	172746	AL645966 Mus musculus chr
gb_pr:GGU08300	+	76.50	122.81	177.35	762	U08300 Gorilla gorilla piron pr	gb_pr:AF418272	-	76.50	70.85	1.4e+05	178952	AF418272 Homo sapiens coa
gb_pr:GGU15166	+	76.50	122.81	177.35	762	U08300 Gorilla gorilla piron pr	gb_hcg:AC012513	+	76.50	70.84	1.4e+05	179149	AC012513 Homo sapiens chr
gb_com:OAAJ738	+	76.50	122.47	185.33	790	AJ000738 Ovis aries PEP gene, a	gb_hcg:AC026331	+	76.50	70.78	1.4e+05	179172	AC026331 Homo sapiens chr
gb_vt:RVU22480	+	76.50	117.37	336.48	1350	U22480 Rabies virus 91350MA nu	gb_hcg:AC027290	+	76.50	70.52	1.4e+05	180339	AC027290 Homo sapiens chr
gb_vt:RVU22481	+	76.50	117.37	336.48	1350	U22480 Rabies virus 8706ARS nu	gb_pr:AP002797	+	76.50	70.31	1.4e+05	180339	AP002797 Homo sapiens chr
gb_vt:RVU22633	+	76.50	117.37	336.48	1350	U22633 Rabies virus 8721ARF nu	gb_hcg:AC069489	-	76.50	70.41	1.4e+05	187388	AC069489 Rattus norvegicus chr
gb_pr:AFV67478	+	76.50	117.16	335.85	1379	AF166478 Homo sapiens annexin	gb_hcg:AL645807	-	76.50	69.55	1.6e+05	189755	AL645807 Mus musculus chr
gb_pr:AY063833	+	76.50	115.38	459.82	1663	AY063833 Arabidopsis thaliana	gb_hcg:AC025558	+	76.50	70.29	1.5e+05	189748	AC025558 Homo sapiens chr
gb_pr:HMMR329	+	76.50	115.35	461.84	1669	U48477 Human skeletal muscle 1	gb_hcg:AL607030	-	76.50	70.29	1.5e+05	189748	AL607030 Mus musculus chr
gb_ba:SNYNEUR	+	76.50	114.61	507.50	1803	M55342 S. typhimurium neuraminid	gb_hcg:AC021155	-	76.50	69.93	1.5e+05	197031	AC021155 Homo sapiens chr
gb_ro:MM082441	+	76.50	113.49	521.52	1844	U82441 Mus musculus EBF-2 mRNA	gb_in:CMS07EG9	-	76.50	69.92	1.5e+05	197466	AL590442 Chromosome II of
gb_pr:AF428342	+	76.50	113.60	549.73	1925	AF428342 Arabidopsis thaliana	gb_in:CMS07EG9	-	76.50	69.92	1.5e+05	197466	AL590442 Chromosome II of
gb_pr:AF428352	+	76.50	113.64	575.28	1958	AF367352 Arabidopsis thaliana	gb_hcg:AL645807	-	76.50	69.55	1.6e+05	205507	AL645807 Mus musculus chr
gb_ro:AF42480	+	76.50	112.94	629.15	2150	AF142480 Rattus norvegicus red	gb_hcg:AC067917	-	76.50	69.23	1.7e+05	205507	AC067917 Homo sapiens chr
gb_ba:BSU80930	+	76.50	112.32	680.97	2294	K00563 b. subtilis alpha-amylase	gb_ba:BSU80020	-	76.50	69.23	1.7e+05	221250	299123 Bacillus subtilis
gb_pr:HSU80930	+	76.50	111.17	789.71	2590	U80930 Human stem cell factor	gb_ba:BSU80020	-	76.50	69.23	1.7e+05	221250	299123 Bacillus subtilis
gb_pr:STNPLMYL	+	76.50	111.03	803.88	2628	K00562 b. subtilis alpha-amylase	gb_pr:AL539920	-	76.50	68.84	1.8e+05	226860	AL539920 Human DNA sequen
gb_ro:MM071189	+	76.50	110.97	809.48	2643	U71189 Mus musculus met-mesenc	gb_hcg:AC096350	-	76.50	68.22	1.9e+05	235932	AC096350 Rattus norvegicus
gb_ba:AF105339	+	76.50	110.94	812.85	2652	AF105339 Neisseria meningitidis	gb_hcg:AC020885	-	76.50	67.00	2.2e+05	268294	AC020885 Mus musculus clo
gb_ba:BSMAYL	+	76.50	110.84	833.34	2680	V00101 Bacillus subtilis amyA	gb_hcg:AC073671	-	76.50	66.69	2.5e+05	276938	AC073671 Mus musculus act
gb_ba:AF012211	+	76.50	109.83	937.20	2980	AF012211 Neisseria meningitidis	gb_ba:TC0104	-	76.50	66.00	2.5e+05	297750	AL445066 Thermoplasma act
gb_ro:AF116519S1	-	76.50	109.58	967.23	3058	AF116519 Mus musculus periplak	gb_ba:AP003129	+	76.50	66.00	2.5e+05	321128	AP003129 Staphylococcus a
gb_pr:PKLPER2	+	76.50	108.82	1.1e+03	3315	U98501 Cloning vector pKIL PCR	gb_hcg:AC078979	+	76.50	63.29	2.7e+05	321128	AC078979 Homo sapiens chr
gb_ba:PM18358	+	76.50	108.81	1.1e+03	3318	U18558 Neisseria meningitidis	gb_ba:AP003358	+	76.50	63.29	2.7e+05	321128	AC078979 Homo sapiens chr
gb_pr:1857939	+	76.50	108.81	1.1e+03	3318	U18558 Neisseria meningitidis	gb_ba:AP003358	+	76.50	63.29	2.7e+05	321128	AC078979 Homo sapiens chr
gb_ba:MM021714	+	76.50	108.30	3499	AB022174 Aeromonas caviae gene	U92729 Sequence 1 from patent	gb_com:AF090852	-	76.00	126.04	117.24	495	AJ227968 Klebsiella oxytoca
gb_ro:MM092703	+	76.50	107.60	1.2e+03	3768	U92729 Sequence 1 from patent	gb_ba:AF090852	-	76.00	126.04	117.24	495	AJ227968 Klebsiella oxytoca
gb_ba:RSMGPN	+	76.50	107.27	1.3e+03	3900	U250182 R. sphaeroides mgs gene	gb_com:AF117312	-	76.00	123.00	173.07	681	AF117312 Equus caballus t1st
gb_in:DDU72236	+	76.50	106.02	1.5e+03	4446	U72236 Dictyostellium mgs gene	gb_com:AF117312	-	76.00	123.00	173.07	681	AF117312 Equus caballus t1st
gb_pat:AF139771	+	76.50	103.39	2.1e+03	5864	ARI39971 Sequence 47 from pat	gb_com:AF117312	-	76.00	123.00	173.07	681	AF117312 Equus caballus t1st
gb_pat:AF140290	+	76.50	103.39	2.1e+03	5864	ARI40290 Sequence 47 from pat	gb_pr:CA075384	+	76.00	122.55	183.36	714	U73384 Cercopithecus aethiops
gb_ba:AF140568	+	76.50	102.68	2.3e+03	5864	ARI40568 Sequence 47 from pat	gb_pr:MS075382	+	76.00	122.55	183.36	714	U73384 Cercopithecus aethiops
gb_ba:BCSP00FA	+	76.50	102.68	2.3e+03	6313	ARI40568 Sequence 47 from pat	gb_pr:AT008812	+	76.00	122.51	184.30	717	U082393 Aotus trivirgatus pri
gb_sy:CVU46197	+	76.50	101.73	2.6e+03	6882	U46197 Cloning vector pCG1662	gb_pr:CM008312	+	76.00	122.43	186.18	723	U083132 Callipepla mollach pri
gb_sy:CVU46197	+	76.50	100.92	2.6e+03	7602	U46197 Cloning vector pCG1662	gb_pr:MS008303	+	76.00	122.24	190.91	738	U083033 Mandrillus sphinx pri
gb_ba:AE006724	+	76.50	98.28	4.1e+03	10029	AE006724 Sulfolobus solfatar	gb_pr:CA008291	+	76.00	122.24	190.91	738	U08291 Cercopithecus aethiop
gb_ba:AE008295	+	76.50	98.16	4.1e+03	10129	AE008295 Agrobacterium tumef	gb_pr:CG075386	+	76.00	122.24	190.91	738	U08292 Cercopithecus aethiop
gb_sy:CVU46196	+	76.50	97.91	4.3e+03	10156	AE01696 Cloning vector pDG1661	gb_pr:CG075386	+	76.00	122.24	190.91	738	U08292 Cercopithecus aethiop
gb_pat:AX100838	+	76.50	97.81	4.4e+03	10426	AX100838 Sequence 82 from pat	gb_pr:CM075387	+	76.00	122.24	190.91	738	U08292 Cercopithecus aethiop
gb_ba:AE004629	+	76.50	97.60	4.5e+03	10541	AE004629 pseudomonas aerugin	gb_pr:CM075387	+	76.00	122.24	190.91	738	U08292 Cercopithecus aethiop
gb_sy:CVU46198	+	76.50	97.60	4.5e+03	10776	U46198 Cloning vector pDG1726	gb_pr:SS015165	+	76.00	121.97	197.56	759	U15165 Saimiri sciureus majo
gb_ba:AE009319	+	76.50	96.97	4.8e+03	11355	AE009319 Agrobacterium tumef	gb_pr:CG008297	+	76.00	121.93	198.52	762	U08297 Colobus guereza pion
gb_in:CEH19M07	+	76.50	96.97	4.9e+03	11509	AE004181 Vibrio cholerae chrc	gb_pr:PM008311	+	76.00	121.93	198.52	762	U08311 Macaca fascicularis pri
gb_ba:RLHUPREG	+	76.50	95.60	5.1e+03	11966	Z95782 B. subtilis chromosome1	gb_pr:PM008311	+	76.00	121.93	198.52	762	U08311 Macaca fascicularis pri
gb_ba:RLHUPREG	+	76.50	95.60	5.1e+03	12766	Z95782 B. subtilis chromosome1	gb_pr:PM008311	+	76.00	121.93	198.52	762	U08311 Macaca fascicularis pri
gb_ba:RLHUPREG	+	76.50	92.83	8.3e+03	17786	X59974 R. leguminosarum bv. v3	gb_pr:MM008307	+	76.00	121.93	198.52	762	U08307 Macaca mulatta pion
gb_pr:AL391420	+	76.50	92.60	8.5e+03	18216	AL391420 Human DNA sequence 1	gb_pr:MM008307	+	76.00	121.93	198.52	762	U08307 Macaca mulatta pion
gb_sy:AF246703	+	76.50	91.08	1.0e+04	21376	AF246703 Synthetic construct	gb_pr:MM008306	+	76.00	121.93	198.52	762	U08306 Macaca mulatta pri
gb_ba:AE008739	+	76.50	90.56	1.1e+04	22578	AE008739 Salmonella typhimur	gb_pr:PM008302	+	76.00	121.93	198.52	762	U08302 Presbytis francoisi p
gb_ba:BSNMA320D	+	76.50	89.23	1.3e+04	25940	Z45782 B. subtilis chromosome1	gb_pr:PM008302	+	76.00	121.93	198.52	762	U08302 Presbytis francoisi p
gb_in:CHBGA41N04	+	76.50	86.39	1.9e+04	34978	AC084590 Caenorhabditis brig	gb_in:AF117322	+	76.00	121.89	199.47	765	AF117322 Rattus norvegicus p
gb_ba:AOCCZCA361	+	76.50	84.73	2.3e+04	41625	AL049826 Streptomyces coelic	gb_com:MP008852	+	76.00	121.78	202.34	774	U08952 Muscivora putorius prio
gb_hcg:AC102954	+	76.50	81.12	3.7e+04	60864	AC102954 Rattus norvegicus ch	gb_pat:AX135431	+	76.00	121.42	231.99	804	AX135431 Sequence 2 from pat
gb_hcg:AC106124	+	76.50	79.97	4.9e+04	71497	AC106124 Rattus norvegicus c1	gb_ba:CL0MTPB	-	76.00	120.47	257.17	808	AL929245 Sequence 2 from pat
gb_hcg:AL606703	+	76.50	77.97	5.5e+04	84711	AP002067 Arabidopsis thaliana	gb_ro:AY074028	+	76.00	119.91	257.17	942	AY074028 Mus musculus olfact
gb_pl:AP002067	+	76.50	76.42	6.7e+04	98653	AC103038 Rattus norvegicus c1	gb_ba:SCS288128	+	76.00	119.87	258.50	946	AJ288128 Serratia sp. partia
gb_hcg:AC103038	+	76.50	76.23	6.9e+04	101720	AC005006 Homo sapiens clone	gb_com:OAAJ736	+	76.00	119.74	263.18	960	AJ000736 Ovis aries PEP gene
gb_hcg:AC022426	+	76.50	76.01	7.1e+04	104115	AC022426 Homo sapiens chrom	gb_pat:AX135431	+	76.00	119.53	348.76	1209	AX135431 Sequence 2 from pat
gb_pr:AC011469	+	76.50	75.43	7.6e+04	110569	AC011469 Homo sapiens chrom	gb_ba:CL0MTPB	-	76.00	117.34	357.94	1235	AB071000 Clostridium thermace
gb_hcg:AC106226	+	76.50	74.94	8.1e+04	116438	AC106226 Rattus norvegicus c	gb_ro:AB047002	-	76.00	117.08	370.00	1269	AB047002 Rattus norvegicus
gb_pr:AC022440	+	76.50	74.84	8.2e+04	117735	AC022440 Homo sapiens ch	gb_ro:RAVAREL	-	76.00	116.24	412.06	1386	J02720 Rat liver arginase m
gb_pr:AC068274	+	76.50	74.74	8.3e+04	118955	AC068274 Homo sapiens BAC cl	gb_pl:AB039746	+	76.00	115.15	473.45	1553	AB039746 Spirodela punctata
gb_ba:DS0453	+	76.50	72.78	1.1e+05	16191	DS0453 Bacillus subtilis DNA	gb_pl:AF434557	+	76.00	114.30	528.34	1699	AF434557 Brassica oleracea
gb_com:AC091119	+	76.50	72.52	1.1e+05	150116	AC091119 Canis familiaris c1	gb_ro:RAVAREL	-	76.00	114.11	541.27	1733	M20513 Rat p10n-related pr
gb_hcg:AC109808	+	76.50	72.46	1.1e+05	151164	AC109808 Homo sapiens chrom	gb_pat:AY075158	+	76.00	114.11	541.27	1733	M20513 Rat p10n-related pr
gb_hcg:AP003217	+	76.50	72.42	1.1e+05	151677	AP003217 Oryza sativa chrom	gb_ba:SGNHR	+	76.00	110.29	883.56	2054	AL1560 Endoxyalanase A (xln
gb_pr:AC023118	+	76.50	72.26	1.1e+05	154330	AC023118 Homo sapiens BAC cl	gb_vt:AY034455	+	76.00	109.67	956.58	2763	AY034455 Equine herpesvirus
gb_hcg:AC006513	+	76.50	71.88	1.2e+05	160574	AC006513 Homo sapiens chrom	gb_pat:AX037558	+	76.00	109.33	999.87	2865	AX037558 Sequence 13 from p
gb_pr:AC079965	+	76.50	71.79	1.2e+05	163494	AC079965 Homo sapiens 12q B7	gb_pat:DS009252	+	76.00	108.61	1.1e+06	3090	DS009252 Rattus norvegicus DN
gb_hcg:AC031978	+	76.50	71.41	1.3e+05	167359	AC031978 Homo sapiens chrom	gb_pat:AX037559	+	76.00	108.56	1.1e+06	3104	AX037559 Sequence 13 from p

gb_vl:D88733	76.00	108.36	1.1e+03	3170	D88733 Equine herpesvirus 1 DN	gb_pr:PTU15039	75.50	121.05	222.21	762	U15039 Pan troglodytes major
gb_da:SHPRP	76.00	105.56	1.6e+03	4256	X01313 Ovis aries prion protei	gb_pr:SSU08308	75.50	121.05	222.21	762	U08308 Symphalangus syndactyl
gb_da:NGSGAPRO	76.00	104.22	1.9e+03	4889	M04835 Neisseria gonorrhoeae g	em_hqo:Inv:AC048110	75.50	118.21	319.89	1077	AC048110 Giardia intestinal
gb_pat:AL12416	76.00	104.22	1.9e+03	4889	A02796 N.gonorrhoeae gene for	gb_pr:HSNBKGCNE	75.50	116.60	399.55	1217	AC098616 H.sapiens mRNA for N
gb_pat:AE0079202	76.00	104.22	1.9e+03	4889	A12416 Iga-protease gene p1cu	gb_pr:HSYPAX71	75.50	114.78	496.42	1472	X96744 H.sapiens PAX7 gene,
gb_da:WSPCRAC	76.00	104.22	1.9e+03	4889	AR07202 Sequence 5 from paten	gb_in:AL1921997	75.50	110.80	827.89	2238	AY061199 Drosophila melanog
gb_da:AF055999	76.00	99.91	3.3e+03	4966	X65042 W.succinogenes genes ps	gb_in:SL123471	75.50	110.75	833.31	2250	AF124371 Spodoptera littura
gb_da:NGCOMCN	76.00	99.09	3.7e+03	8333	Z49855 N.gonorrhoeae gene for	gb_ro:AF098460	75.50	108.73	1.1e+03	2780	AF098460 Mus musculus Link
gb_hhg:AC109577	76.00	98.42	4.0e+03	9007	AC109577 Rattus norvegicus c1c	gb_sy:PKLIPCL	75.50	107.31	1.3e+03	3228	X98500 Cloning vector PKL
gb_pr:AE144094	76.00	95.79	5.7e+03	11876	AF144094 Homo sapiens unconv	gb_da:KPSRERB	75.50	106.08	2.0e+03	3675	M13459 Serratia serine prot
gb_da:AE000897	76.00	92.97	8.1e+03	12027	M80537 Drosophila melanogaste	gb_pr:AF291999	75.50	103.90	2.0e+03	4621	X57401 R.pneumoniae scry, s
gb_in:IDROFAR	76.00	92.97	8.1e+03	15066	M80537 Drosophila melanogaste	gb_da:IPORGES	75.50	101.68	2.7e+03	5835	AF291999 Pneumocystis carin
gb_in:CHRG39112	76.00	85.03	2.3e+04	36774	AC087161 Leishmania major chr	gb_da:AE009481	75.50	96.69	5.1e+03	9850	X62540 P.putida genes rpnH,
gb_hhg:AC109915	76.00	83.99	2.6e+04	41040	AC084573 Caenorhabditis br19g	gb_da:AE000248	75.50	95.91	5.5e+03	10683	AE009481 Brucella melitens
gb_hhg:AC109915	76.00	79.71	4.4e+04	64306	AC109515 Homo sapiens chr19g	gb_da:AE0004614	75.50	95.64	5.8e+03	10998	AE004614 Pseudomonas aerug
gb_pr:AL445429	76.00	78.92	4.9e+04	69926	AC109315 Homo sapiens chr19g	gb_da:AE0005407	75.50	95.51	5.9e+03	11155	AE005407 Escherichia coli1
gb_in:AC004573	76.00	77.05	6.2e+04	85095	AC004573 Drosophila melanogae	gb_da:AE009504	75.50	94.93	6.3e+03	11851	AE009504 Brucella melitens
gb_pr:AL512407	76.00	76.65	6.5e+04	88376	AL512407 Human DNA sequence f	gb_da:AE009827	75.50	90.50	1.1e+04	18886	D90827 E.coli genomic DNA,
gb_hhg:AC018551	76.00	75.46	6.7e+04	90477	AC018551 Homo sapiens chr19g	gb_da:AF372703	75.50	89.60	1.3e+04	20741	AF372703 Pseudomonas syrin
gb_pr:AP005196	76.00	75.97	7.1e+04	95555	AF014976 Homo sapiens chr19g	gb_da:ALM243431	75.50	87.11	1.7e+04	26553	AF243431 Actinobacter two
gb_pr:AP000128	76.00	75.51	7.6e+04	100000	AP000128 Homo sapiens unconv	gb_in:CEB0391	75.50	86.18	1.9e+04	29731	B8154 Caenorhabditis eleg
gb_pr:AP000206	76.00	75.51	7.6e+04	100000	AP000206 Homo sapiens genom	gb_hhg:AC098373	75.50	85.29	2.2e+04	32621	AF298262 Porcine adenoviru
gb_pr:AT118012	76.00	75.45	7.6e+04	100678	AL138644 Arabidopsis thalian	gb_in:CHRG11A11	75.50	82.95	2.8e+04	39784	AC084373 Rattus norvegicus
gb_hhg:AC073702_0	76.00	74.43	8.5e+04	110000	AC073702 Mus musculus clone	gb_hhg:AC023620	75.50	82.36	3.2e+04	44411	AC084456 Caenorhabditis br
gb_da:ID90899	76.00	72.74	1.1e+05	133859	D90899 Synchocystis sp. PC	gb_hhg:AC083369	75.50	80.79	3.9e+04	52352	AC083369 Homo sapiens chro
gb_pr:AL157709	76.00	72.73	1.1e+05	133865	AL157709 Human DNA sequenc	gb_in:AC006402	75.50	77.50	3.9e+04	73544	AC006402 Drosophila melano
gb_vl:AF147806	76.00	72.40	1.1e+05	138675	AF147806 Gallid herpesvirus	gb_pr:AB010074	75.50	76.04	7.1e+04	86212	AB010074 Arabidopsis thail
gb_hhg:AC01491	76.00	72.16	1.2e+05	142257	AC01491 Drosophila melanog	gb_hhg:AC095700	75.50	75.96	7.1e+04	86645	AC095700 Rattus norvegicus
gb_pr:AC012331	76.00	72.08	1.2e+05	143376	AC012331 Homo sapiens, comp	gb_pr:ATP9N11	75.50	75.57	7.2e+04	78635	AL109756 Arabidopsis thail
gb_hhg:OSJN00091	76.00	71.89	1.2e+05	146255	AL606646 Oryza sativa chrom	gb_hhg:AC095887	75.50	73.69	9.5e+04	110409	AC095887 Rattus norvegicu
gb_hhg:AC023916	76.00	71.73	1.2e+05	146389	AC023916 Homo sapiens chr19g	gb_pr:AC002477	75.50	72.98	1.1e+05	124095	AC002477 Human PAC clone
gb_pl:AP003725	76.00	71.73	1.2e+05	146706	AP003725 Oryza sativa genom	gb_hhg:AC005677	75.50	71.94	1.2e+05	136353	AC005677 Drosophila melan
gb_hhg:AC105683	76.00	71.70	1.2e+05	149212	AC105683 Rattus norvegicus c	gb_pl:AP003854	75.50	71.62	1.2e+05	134578	AP003854 Oryza sativa gen
gb_vl:HSBSCOM8	76.00	71.64	1.2e+05	150223	M86664 Equine herpesvirus 1	gb_hhg:AC003932	75.50	71.80	1.2e+05	137259	AC003932 Homo sapiens chr
gb_pr:AC096588	76.00	71.61	1.2e+05	150864	AC096588 Homo sapiens chr19g	gb_hhg:AC003300	75.50	71.35	1.2e+05	137878	AP003300 Oryza sativa chr
gb_pr:AC002308	76.00	71.58	1.2e+05	151111	AC002308 Homo sapiens chr19g	gb_pr:AL193197	75.50	71.35	1.3e+05	141124	AL193197 Human DNA sequen
gb_hhg:AC010802	76.00	71.52	1.3e+05	152009	AC010802 Homo sapiens chr19g	gb_pr:AC005686	75.50	71.34	1.3e+05	141387	AC005686 Homo sapiens clo
gb_pr:AC068616	76.00	71.49	1.3e+05	153573	AC068616 Homo sapiens chr19g	gb_hhg:AC067789	75.50	71.25	1.3e+05	146616	AC067789 Homo sapiens clo
gb_pr:AP000245	76.00	71.04	1.3e+05	155912	AP000245 Homo sapiens BAC c1	gb_hhg:AL1931624	75.50	70.97	1.3e+05	146871	AL1931624 Human DNA sequen
gb_hhg:AC068048	76.00	70.95	1.3e+05	161485	AC068048 Homo sapiens chr19g	gb_pr:AC005913	75.50	70.96	1.3e+05	147124	AC005913 Homo sapiens xp2
gb_hhg:AC087164	76.00	70.74	1.4e+05	165090	AC087164 Homo sapiens chr19g	gb_hhg:AL1589985	75.50	70.80	1.4e+05	149490	AL158985 Homo sapiens chr
gb_hhg:AC061740	76.00	70.62	1.4e+05	167106	AC061740 Homo sapiens chr19g	gb_hhg:AC010228	75.50	70.72	1.4e+05	150803	AC010228 Homo sapiens chr
gb_hhg:AC061950	76.00	70.52	1.4e+05	167503	AC061950 Homo sapiens chr19g	gb_hhg:AC018409	75.50	70.49	1.4e+05	154468	AC018409 Homo sapiens clo
gb_hhg:AL358216	76.00	70.52	1.4e+05	168995	AL358216 Homo sapiens chr19g	gb_hhg:AC010942	75.50	70.49	1.4e+05	154523	AC010942 Homo sapiens chr
gb_hhg:AC087164	76.00	70.38	1.4e+05	171415	AP002771 Homo sapiens chr19g	gb_hhg:AC020006	75.50	70.34	1.5e+05	157749	AC020006 Drosophila melan
gb_pr:AC006376	76.00	70.35	1.5e+05	172007	AC006376 Homo sapiens BAC c1	gb_hhg:AC025535	75.50	70.29	1.5e+05	157749	AC025535 Homo sapiens chr
gb_hhg:AC108659	76.00	70.01	1.5e+05	180379	AC108659 Homo sapiens 12 BAC	gb_pr:AC007003	75.50	70.27	1.5e+05	157756	AC007003 Homo sapiens BAC
gb_hhg:AC106455	76.00	69.90	1.5e+05	180379	AC106455 Rattus norvegicus c	gb_pl:AP003492	75.50	70.27	1.5e+05	158084	AP003492 Oryza sativa gen
gb_hhg:AC084101	76.00	69.87	1.5e+05	180920	AC084101 Homo sapiens chr19g	gb_hhg:AC0093306	75.50	70.26	1.5e+05	158231	AC0093306 Homo sapiens chr
gb_hhg:AC104580	76.00	69.78	1.6e+05	186511	AC104580 Homo sapiens chr19g	gb_ro:AL133491	75.50	70.26	1.5e+05	158344	AL133491 Mouse DNA sequen
gb_in:AC008002	76.00	69.77	1.6e+05	187262	AC108002 Drosophila melanog	gb_hhg:AC025997	75.50	70.25	1.5e+05	158523	AC025997 Homo sapiens chr
gb_hhg:AC009713	76.00	69.24	1.7e+05	193196	AC009713 Mus musculus, clone	gb_hhg:AC016621	75.50	70.19	1.5e+05	159438	AC016621 Homo sapiens chr
gb_ro:AC021067	76.00	68.72	1.8e+05	203101	AC021067 Mus musculus, clone	gb_hhg:AC009951	75.50	70.12	1.5e+05	160669	AC009951 Homo sapiens chr
gb_pr:AC016397	76.00	68.59	1.8e+05	206903	AC016397 Homo sapiens chr19g	gb_hhg:AC022920	75.50	70.12	1.5e+05	160676	AC022920 Homo sapiens chr
gb_hhg:AC046168	76.00	68.45	1.8e+05	209788	AC046168 Homo sapiens chr19g	gb_pr:AC021517	75.50	70.10	1.5e+05	160904	AC021517 Homo sapiens chr
gb_hhg:AC009782	76.00	68.45	1.8e+05	210036	AC009782 Homo sapiens chr19g	gb_hhg:AC021212	75.50	70.06	1.5e+05	161584	AC021212 Homo sapiens chr
gb_da:AE001862	76.00	68.28	1.9e+05	213732	AE001862 Deinococcus radiod	gb_hhg:AC026106	75.50	69.89	1.5e+05	161420	AC026106 Homo sapiens chr
gb_hhg:AL307078	76.00	67.95	2.0e+05	221372	AL307078 Homo sapiens chr19g	gb_hhg:AC096023	75.50	69.69	1.6e+05	166125	AC096023 Rattus norvegicu
gb_hhg:AC090434	76.00	67.08	2.2e+05	244093	AC090434 Chlamydomonas reinb	gb_hhg:AC009951	75.50	69.46	1.6e+05	172151	AC009951 Homo sapiens chr
gb_hhg:AC023338	76.00	66.86	2.2e+05	244093	AC023338 Homo sapiens chr19g	gb_hhg:AC017058	75.50	69.46	1.6e+05	172151	AC017058 Homo sapiens chr
gb_da:AL627273	76.00	66.79	2.3e+05	250500	AL627273 Salmonella enterica	gb_pr:AC011841	75.50	69.33	1.7e+05	174588	AC011841 Homo sapiens chr
gb_hhg:AC096620	76.00	66.46	2.4e+05	258051	AC096620 Rattus norvegicus c	gb_hhg:AC0068507	75.50	69.24	1.7e+05	176282	AC068507 Homo sapiens chr
gb_hhg:AC105485	76.00	66.46	2.4e+05	258779	AC105485 Rattus norvegicus c	gb_hhg:AC015714	75.50	69.23	1.7e+05	176307	AC015714 Homo sapiens chr
gb_hhg:AC0021061	76.00	65.29	2.6e+05	284664	AC0021061 Mus musculus chr19g	gb_hhg:AC025544	75.50	69.21	1.7e+05	176728	AC025544 Homo sapiens chr
gb_da:AP001513	76.00	64.98	2.8e+05	302473	AP001513 Bacillus halodurans	gb_hhg:AL450083	75.50	69.19	1.7e+05	176914	AL450083 Homo sapiens chr
gb_pr:AP003589	76.00	64.98	2.8e+05	302473	AP003589 Drosophila melanog	gb_hhg:AC0097891	75.50	69.19	1.7e+05	177206	AC0097891 Rattus norvegicu
gb_da:AP003001	76.00	63.86	3.2e+05	340000	AP003001 Homo sapiens genom	gb_hhg:AC094889	75.50	69.12	1.7e+05	177913	AC094889 Rattus norvegicu
gb_pr:HU082829	75.50	63.70	3.3e+05	347873	AP003001 Mesorhizobium loti	gb_hhg:AC067897	75.50	69.09	1.7e+05	178951	AC067897 Homo sapiens chr
gb_pr:PT008236	75.50	121.05	222.21	762	U08296 Pan troglodytes prion pr	gb_hhg:AC108703	75.50	69.09	1.7e+05	178967	AC108703 Homo sapiens chr


```

gb_ba:AF065404 + 75.50 68.95 1.7e+05 181654 ! AF065404 Bacillus anthracis
gb_pr:AC079239 + 75.50 68.91 1.7e+05 182406 ! AC079239 Homo sapiens BAC C1
gb_in:AC093198 + 75.50 68.88 1.7e+05 182901 ! AC093198 Drosophila melanoga
gb_ba:AP000063 - 75.50 68.76 1.8e+05 185300 ! AP000063 Aeropyrum pernix ge
gb_hhg:AC079560 + 75.50 68.60 1.8e+05 188406 ! AC079560 Mus musculus clone
gb_pr:AC022113 - 75.50 68.59 1.8e+05 188570 ! AC022113 Homo sapiens chromo
gb_hhg:AC073658 + 75.50 68.58 1.8e+05 188873 ! AC073658 Mus musculus clone
gb_in:AC093440 + 75.50 68.42 1.9e+05 192132 ! AC093440 Drosophila melanoga
gb_hhg:AC025784 + 75.50 68.33 1.9e+05 193961 ! AC025784 Mus musculus clone
gb_pr:AL358975 - 75.50 68.21 1.9e+05 196337 ! AL358975 Human DNA sequence
gb_hhg:AC013569 + 75.50 68.12 1.9e+05 198160 ! AC013569 Homo sapiens clone

```

seq_name: gb_ba:AF242417

seq_documentation_block:

LOCUS AF242417 777 bp DNA linear BCT 01-JAN-2002
DEFINITION Escherichia coli heat-labile enterotoxin subunit A precursor, gene,
complete cds.

ACCESSION AF242417 GI:18026885

VERSION AF242417

KEYWORDS

SOURCE

ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE 1 (bases 1 to 777)

AUTHORS Chang, G.N. and Ho, K.C.

TITLE Heat-labile enterotoxin subunit A gene of Escherichia coli

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Chang, G.N. and Ho, K.C.

TITLE Direct Submission

JOURNAL Submitted (07-MAR-2000) Botany, National Taiwan University, 1

Roosevelt Road, Sec. 4, Taipei, Taiwan 106, Republic of China

FEATURES Location/Qualifiers

SOURCE

1..777
/organism="Escherichia coli"

/strain="CH-C2"

/serotype="O-136"

/specific_host="pig"

/db_xref="taxon:562"

/country="Taiwan"

/note="Isolated from hemolytic cells"

CDS

1..777

/note="LtpA"

/codon_start=1

/product="heat-labile enterotoxin subunit A precursor"

/protein_id="AA155671.1"

/db_xref="GI:18026886"

/translation="MKNTFFIFILLASPLVANGDRLYRADSRPDEIKRSGGLMPRG

HNEVFDRTGOMNINLDHARQTGFVRDDVSTLSLRSAHLACGSLISGYSTY

IYVIAFPMFVNVDLVGYSPHPEOEVSADGGIPIYSIYGVNFGVIDERLHN

REYRDYRYNNINIAPEADGYRLAGRPDHOAREPPIHHAPGGGNSRTTGTCTN

ETTONLSTYLRYSKYKQIFSDYQSEVDYNNIRDEL"

sig_peptide

mat_peptide

mat_peptide

55..630
/product="heat-labile enterotoxin subunit A1"

637..774

/product="heat-labile enterotoxin subunit A2"

BASE COUNT

255 a

136 c

164 g

222 t

ORIGIN

alignment_scores:

Quality: 1304.00

Ratio: 5.433

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AF242417

Align seg 1/1 to: AF242417 from: 1 to: 777

seq_name: gb_ba:P307LTA

seq_documentation_block:

LOCUS P307LTA 777 bp DNA linear BCT 24-APR-1996

DEFINITION Plasmid P307 (from E. coli) heat-labile enterotoxin subunit A (lta)

ACCESSION M35581

VERSION M35581.1 GI:150458

KEYWORDS enterotoxin; lta gene.

SOURCE Plasmid P307 (clone: pAT151.) DNA.

ORGANISM Plasmids.

Plasmids.

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluLeu 17
|||||
55 AATGGGACAGATATACCGTGTGACTGACAGCCCGAGATGAATATA 104
|||||
17 sAtSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACCTTCGGGAGGCTTATGCCAGAGGCAATATGAGTCTCGATGAG 154
|||||
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCTAATGAATATTTATCTTATGATCACCAGGAGACCAAC 204
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||
205 GCGTTTGTCGATATGATGACGATATGTTCCACTTCCTCTGTTGAG 254
|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||
255 AAGTGTCACTTAGCAGACAGCTATATTATCAGATATTCACCTTACT 304
|||||
84 YrLLeuTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTTATATAGCAGACACCAATATGTTATATGATGATG 354
|||||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
355 TTAGGGGTATACAGCCCTCACATATGAACAGAGGTTCCTCGTTAG 404
|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAspPheGly 134
|||||
405 TGGAAATACCATTTCTCAGATATATGATGATGCTGCTGTTAAATTTGG 454
|||||
134 aLleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
455 TGAATGATGACCATTTACATCGTAAACAGGGAATATGAGACCGGTTAT 504
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||
505 AGAATCTGAATATAGCTCCGACAGAGATGTTACAGATTACAGAGTT 554
|||||
167 eProProAspHisGlnAlaLeuPargGluGluProTPIHisHisAla 184
|||||
555 CCCACCGATCACCAAGCTTGAGAGAAACCTGGATCATCATGAC 604
|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAGGTTTGGAATTCATCAAGAACAATCACAGGTATATTGTAAT 654
|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
655 GAGGAGACCCAGATCTGACACAAATATATCTCAGGAATATCAATCA 704
|||||
217 sValTyrArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
705 AGTTAAGGACAGATATTTTCAGACTATCTGTCAGAGGTTGACATATA 754
|||||
234 snArgIleArgAspGluLeu 240
|||||
755 ACAGAAATTCGAGATGAATTA 774
|||||

```


REFERENCE 1 (bases 1 to 777)
 AUTHORS Dykes, C.W., Halliday, I.J., Hobden, A.N., Read, M.J. and Harford, S.
 TITLE A comparison of the nucleotide sequence of the A subunit of heat-labile enterotoxin and cholera toxin
 JOURNAL FEMS Microbiol. Lett. 26, 171-174 (1985)
 FEATURES
 source 1. 777
 /organism="Plasmid P307"
 /plasmid="plasmid P307"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2472"
 /clone="PAT153."
 1. 777
 /gene="lta"
 1. 777
 /gene="lta"
 1. 777
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AAA98202.1"
 /db_xref="GI:150459"

CDS
 gene

BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x P307LTA ..

Align seg 1/1 to: P307LTA from: 1 to: 777

```

1  AaGgIyAsPaRgLeuTyTaRgAlaAsPSeTaRgProRoAsPgluIlely 17
   |||||||
55  AATGGCGCACGATTAATACCGTCTGACTCTAGACCCCAAGATCAATAAA 104
   |||||||
17  sAtgSeRgIyLeuMeTProARgIyHIsAnGluTyTPhAsPrg 34
   |||||||
105 ACGTTCGGAGGTCTTATGCCAGAGGCGCAATAGACTCTTGCATGAG 154
   |||||||
34  lYtHrGlMeTAsnIleAsnLeuTyTaRgPHisAlaARgIyThrGlntHr 50
   |||||||
155 GAACCTCAATGAAATTAATCTTATGATCAGCGAGAGGAGACACAACC 204
   |||||||
51  GlYPHeValARgTyTaRgAsPaRgIyTyTaRgSerThrSerLeuSerLeuAR 67
   |||||||
205 GCGTTTGTCTGATATGATGATGATGATGATGATGATGATGATGATGAG 254
   |||||||
67  gSeRAlaHIsleuAlaGluInSerIleleuSeRgIyTyTaRgSerThrTyT 84
   |||||||
255 AAGGACACACTAGACAGACAGTCTATATATATCAGATATTCACCTTACT 304
   |||||||
84  yIleTyValIleAlaThRaLProAsnMeTPhAsnValAsnAsPVal 100
   |||||||
305 ATATATATGTATATAGCGACAGCAACAAATATGTTTAAATGATGATGA 354
   |||||||
101 LeuGlyValIySeRProHIsProTyRgIuGluInIuValSeRAlaLeuG 117
   |||||||
355 TTAGGCGCTATACAGCCCTCACCATATAGACAGAGAGGTTCTGCGTTAG 404
   |||||||
117 yGyIleRToTySeRGlInIleTyRgIyTPrTyTaRgValAsnRheGly 134
   |||||||
405 TGGAAATCCATATCTCAGATATATGATGATGATGATGATGATGATGAT 454
   |||||||
134 AlleAsRgIuAlRgLeuHIsARgAsnARgIuTyTaRgARgATgTyTyT 150

```

seq_name: gb_pat:A04913

seq_documentation_block:

LOCUS A04913 777 bp DNA linear PAT 13-JUL-1993
 DEFINITION E.coli LTA gene for labile toxin A.
 ACCESSION A04913
 VERSION A04913.1 GI:412215
 KEYWORDS labile toxin A.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes, M.V., Harford, S. and Ross, G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 5 19-JUN-1985;
 GLAXO GROUP LIMITED

FEATURES
 location/Qualifiers

source

CDS

1. 777
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1. 777
 /codon_start=1
 /transl_table=11
 /product="labile toxin A"
 /protein_id="CAA00402.1"
 /db_xref="GI:412216"
 /db_xref="SWISS-PROT:P06717"
 /translation="MKNTFFIFFILLASPLYANGDRLYRADSRPPDEIKRSGGLMPRG
 HNEYFDRTQMNTINLYDHARGTGTGFVRYDGYSTSLRSANHLAAGSILSGYSTYY
 IYVATAPNMFVNVDVLGVYSPHPYEOEVSALGIPYQIYGMVYVNFVGLDERLHRN
 REYRDYRYRNINIAFAEDGYRLAGFPDPHOAMREPRWIHNAFOGCGNSSRITGDTCN
 EETONLSRTIYREYOSKYVKROIIFS DYOSEVDIYNIRDEL"

BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A04913 ..

Align seg 1/1 to: A04913 from: 1 to: 777


```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
|||||
55 AATGGGACAGATTATACCGTCTGACTGACGCCCAATATAAATAA 104
|||||
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACCTTCCGGAGGCTTATGCGCCAGAGGCAATATGAGTACTCGATAGAG 154
|||||
34 LyrHrcInMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCTCAATGAAATATTAACTTTATGATCAGCGAGAGAACAAAC 204
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||
205 GCCTTGTGATGATATGATGACGATATGTTCCACTTCTTGTAG 254
|||||
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||
255 AAGTGTCACTTATGACAGACAGCTATATTATCAGATATTCACCTACT 304
|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATTATATGTTATAGCGACAGCACCAATATGTTAATGTTAATGATGA 354
|||||
101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
|||||
355 TTAGGGGTATACAGCCCTCCATATGAAACAGAGGTTCTCTCGTTAG 404
|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
|||||
405 TGGAAATACCATTTCTCAGATATATGATGATGTTATGTTAATTTGGTG 454
|||||
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
455 TGAATTGACGATTTACATCGTACAGGGAATATAGAGCCGGTATAC 504
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||
505 AGAATCTGATATAGTCTCGGACAGAGATGTTACAGATTACAGAGTTT 554
|||||
167 eProProAspHisGlnAlaThrArgGluGluProTyrPheHisAlaP 184
|||||
555 CCCACCGGATACCAAGCTTGAGAGAGAACCTGATTCATCATCAGCAC 604
|||||
184 rGlnGlyCysGlyAsnSerSerArgThrIleHrgIleAspThrCysAsn 200
|||||
605 CACAGGTTGTGAAATTCATCAAGAACATCAACAGGTGATCTGTAAT 654
|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
655 GAGGAGACCCAGATCTGAGCACATATATCTCAGGGAATATCAATCAA 704
|||||
217 sValIleValArgGlnIlePheSerAspTyrGlnSerGluValAspIle 234
|||||
705 AGTTAAGAGCAGATATTTTCAGACATATCAGTACAGAGTTGACATATA 754
|||||
234 snArgIleArgAspGluLeu 240
|||||
755 ACAGATTCGGGATGAATTA 774
|||||
seq_name: gb_ba:ENSLTPA
seq_documentation_block:
LOCUS ENSLTPA 866 bp DNA linear BCT-26-APR-1993
DEFINITION Plasmid ENT P307 (enterotoxigenic E.coli) LTP subunit A gene.
ACCESSION M15362 GI:148337
VERSION M15362.1
KEYWORDS LTP gene; heat-labile toxin.
SOURCE Plasmid ENT (enterotoxigenic E.coli porcine isolate) DNA, clone
P370.
ORGANISM plasmid ENT
plasmids.

```

```

REFERENCE 1 (bases 1 to 866)
AUTHORS Yamamoto,T., Gojobori,T. and Yokota,T.
TITLE Evolutionary origin of pathogenic determinants in enterotoxigenic
JOURNAL Escherichia coli and Vibrio cholerae O1
MEDLINE J. Bacteriol. 169, 1352-1357 (1987)
87137303
COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto
(22-APR-1987).
FEATURES
source location/Qualifiers
1..866
/organism="Plasmid ENT"
/plasmid="Plasmid ENT"
/db_xref="taxon:2463"
90..143
/note="Heat-labile toxin subunit A signal peptide"
90..866
/note="Heat-labile toxin subunit A precursor"
CDS
sig_peptide
90..143
/codon_start=1
/transl_table=1
/protein_id="AA024793.1"
/db_xref="GI:148338"
/translation="MKNTFFIFLLASPLYANGDRLRADSRPPDEIKRSGLMPRG
HNEYFDRTQKINIMLYDHAKRTQGFVRYDGVYSTLSASALAGSITSGSYTY
LYVATAPNMFNVDVIGVSPHPPEEVSALGIPYSQITGTRVNFVITDELHNR
REYRDYRNINLIPAEQGYRLAGFPEDHQWREPWIHAPDQCGNSSRTTSDTCN
EETONLSTIYLRVYQSKVKRQIFSDYCEVDIYRIRDEL"
144..863
mat_peptide
144..863
/product="Heat-labile toxin subunit A"
BASE COUNT 277 a 149 c 181 g 259 t
ORIGIN Unreported.
alignment_scores:
Quality: 1304.00 Length: 240
Ratio: 5.433 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x ENSLTPA ..
Align seg 1/1 to: ENSLTPA from: 1 to: 866
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
|||||
144 AATGGGACAGATTATACCGTCTGACTGACGCCCAATATAAATAA 193
|||||
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
194 ACCTTCCGGAGGCTTATGCGCCAGAGGCAATATGAGTACTCGATAGAG 243
|||||
34 LyrHrcInMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
244 GAACCTCAATGAAATATTAACTTTATGATCAGCGAGAGAACAAAC 293
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||
294 GCCTTGTGATGATATGATGACGATATGTTCCACTTCTTGTAG 343
|||||
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||
344 AAGTGTCACTTATGACAGACAGCTATATTATCAGGATTTCCACTTACT 393
|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
394 ATATTATATGTTATAGCGACAGCAAAATATGTTAATGTTATGATGATGA 443
|||||
101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
|||||
444 TTAGGGGTATACAGCCCTCCATATGAAACAGAGGTTCTCTCGTTAG 493
|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
|||||
494 TGGAAATACCATTTCTCAGATATATGATGATGTTATGTTAATTTGGTG 543
|||||

```


134 a1lleaSPgluArgleuH1sArgasnArgluuTyraArgaspArgTyrr 150
 |||||||
 544 TGATGTGATGAACGATTTCATCGTACAGGAAATATAGACCGGATTATAC 593
 |||||||
 151 ArgasnLeuasn1leAlaProAlaGluaspGlyTyraArgleuAlaGlyph 167
 |||||||
 594 AGAATCTGAATATAGCTCCGGCAGAGAGATGTTACAGATTAGAGGTTT 643
 |||||||
 167 ePrProAspHisGlnAlaTrpArgGluGluProTrrPleHisAlaP 184
 |||||||
 644 CCCACCGGATCCACCAACCTTGAGAGAAACCTGGATTCAATCATCATC 693
 |||||||
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysasn 200
 |||||||
 694 CACAGAGTTTGGAATTCATCAAGAACATACACAGGTGATCTTGTAAAT 743
 |||||||
 201 GluGluThrGlnAsnLeuSerThrIleTyrlleuArgGluuTyrglnSer 217
 |||||||
 744 GAGAGACCCAGAAATCTGACACAAATATCTCAGGGAATATCAATCAA 793
 |||||||
 217 sVallyArgGlnIlePheSerAspTyrglnSerGluValAspIleTyra 234
 |||||||
 794 AGTTAAAGCGCAGATATTTTCAGACTATCAGTTCAGAGTTGACATATATA 843
 |||||||
 234 snArgIleArgaspGluLeu 240
 |||||||
 844 ACAGAAATTCGGGATGAATTA 863
 |||||||
 seq_name: gb_ba:ENRLTPA

seq_documentation_block:
 LOCUS ENRLTPA 898 bp DNA linear BCT-26-APR-1993
 DEFINITION Plasmid ENT-R pCC86 (enterotoxigenic E. coli) LTP subunit A gene.
 ACCESSION M15361
 VERSION M15361.1 GI:148333
 KEYWORDS LTP gene; heat-labile toxin.
 SOURCE Plasmid ENT-R (enterotoxigenic E. coli porcine isolate) DNA, clone pCC86.
 ORGANISM Plasmid ENT-R
 plasmids.
 1 (bases 1 to 898)
 Yamamoto, T., Gotohori, T. and Yokota, T.
 Evolutionary origin of pathogenic determinants in enterotoxigenic
 Escherichia coli and vibrio cholerae O1
 J. Bacteriol. 169, 1352-1357 (1987)
 JOURNAL MEDLINE
 COMMENT Clean copy of sequence [1] kindly provided by T. Yamamoto
 (22-Apr-1987).
 FEATURES
 source location/Qualifiers
 1..898
 /organism="Plasmid ENT-R"
 /plasmid="Plasmid ENT-R"
 /db_xref="taxon:2462"
 122..175
 /note="Heat labile toxin subunit A signal peptide"
 122..898
 /note="Heat labile toxin subunit A precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AAA24791.1"
 /db_xref="GI:148334"
 /translation="MKNTFFILLASPIYANGRLYRADSPPEDEIKRSGLPRG
 HNEFYDRSTQMKINILYDHAHQGTQTFVRIDGIVSTSLKSAHLAQSILSGYSTY
 IYIVATAPNMFENVNDLVGVSPHPEQEVSAIGGLPYQIYGVYVNFSGVIDERLHRN
 REYRDYRVNLINIAPEADGVSLAGFPDPHOAMREBPWIIHAFQGGNSRFTTGDTCN
 EETQNLSTIYLYREYSKYVKRQIFSDYQSEVDIYNRIDEL"
 176..895
 /product="Heat labile toxin subunit A"
 BASE COUNT 289 a 155 c 106 g 268 t
 ORIGIN Unreported.

alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x ENRLTPA ..
 Align seq 1/1 to: ENRLTPA from: 1 to: 898

1 AsnGlyAspArgleuTyraArgAlaAspSerArgProProAspGluIleTy 17
 |||||||
 176 AATGGCAGACAGATTAATACCGGTGACTGTAGACCCAGATGAATAAAA 225
 |||||||
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrrPheaspArg 34
 |||||||
 226 ACGTTCGGAGGCTCTATGCCAGAGGCGATTAAGAGTACTTCGATAGAG 275
 |||||||
 34 lyThrGlnMetAsn1leAsnLeuTyraSpHisAlaArgGlyThrGlnThr 50
 |||||||
 276 GAACCTCAAATGAATATTATATCTTTATGATCAGCGAGAGAACCAAC 325
 |||||||
 51 GlyPheValArgTyraAspAspGlyTyraValSerThrSerLeuSerIle 67
 |||||||
 326 GCGTTTGACAGATATGATGAACGGATATGTTCCACTTCTGTACTTGAG 375
 |||||||
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrr 84
 |||||||
 376 AAGTCTCACCTTACAGACAGCTATATATATCAGGATATTCACACTTACT 425
 |||||||
 84 yrlleTyraVal1leAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||||
 426 ATATATATGTTATATAGCCAGACACCAATATGTTAAATGTTAATGATGA 475
 |||||||
 101 LeuGlyValTyrrSerProHisProTyrglnGlnIleValSerAlaLeuG 117
 |||||||
 476 TTAGGCGTATACAGCCCTCACCATATGAACAGAGGTTCTGCGTTAGG 525
 |||||||
 117 yGlyIleProTyrrSerGlnIleTyrglyTyrrTyraValAsnPheGly 134
 |||||||
 526 TGGAAATACCAATATCTCAGATATGTGATGGTATACGTATATTTGGTG 575
 |||||||
 134 a1lleaSPgluArgleuH1sArgasnArgluuTyraArgaspArgTyrr 150
 |||||||
 576 TGATGTGATGAACGATTTCATCGTACAGGAAATATAGACCGGATTATAC 625
 |||||||
 151 ArgasnLeuasn1leAlaProAlaGluaspGlyTyraArgleuAlaGlyph 167
 |||||||
 626 AGAATCTGAATATAGCTCCGGCAGAGAGATGTTACAGATTAGAGGTTT 675
 |||||||
 167 ePrProAspHisGlnAlaTrpArgGluGluProTrrPleHisAlaP 184
 |||||||
 676 CCCACCGGATCCACCAACCTTGAGAGAAACCTGGATTCAATCATCATC 725
 |||||||
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysasn 200
 |||||||
 726 CACAGAGTTTGGAATTCATCAAGAACATACACAGGTGATCTTGTAAAT 775
 |||||||
 201 GluGluThrGlnAsnLeuSerThrIleTyrlleuArgGluuTyrglnSer 217
 |||||||
 776 GAGAGACCCAGAAATCTGACACAAATATCTCAGGGAATATCAATCAA 825
 |||||||
 217 sVallyArgGlnIlePheSerAspTyrglnSerGluValAspIleTyra 234
 |||||||
 826 AGTTAAAGCGCAGATATTTTCAGACTATCAGTTCAGAGTTGACATATATA 875
 |||||||
 234 snArgIleArgaspGluLeu 240
 |||||||
 876 ACAGAAATTCGGGATGAATTA 895
 |||||||
 seq_name: gb_pat:A04915
 seq_documentation_block:

LOCUS A04915 777 bp DNA linear PAT 14-JUL-1993
 DEFINITION Artificial sequence LTA gene for labile toxin A.
 ACCESSION A04915
 VERSION A04915.1 GI:412519
 KEYWORDS labile toxin A.
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes, M.V., Harford, S. and Ross, G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 7 19-JUN-1985;
 GLAXO GROUP LIMITED
 FEATURES
 source Location/Qualifiers
 1..777
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..777
 /codon_start=1
 /transl_table=11
 /product="labile toxin A"
 /protein_id="CAA00403.1"
 /db_xref="GI:412520"
 /translation="MKNITFTFILLASPLVANGDRLYRADSRPPDEIKRSGCLMPRG
 HNEYPDRGTOMNINLYDHARGQTFGEVRYDGYVYTSLSLRSAHLAGOSILSGYSTLY
 IYVIATPMMFVNDVLYGYSBHPYEOEYSEAGLIPYSQIYGMVHNFGVIDERLHRN
 REYRDYRYNLIAPAEDEGYRLAGFPDHOAWREBPWHIHAQGGGNSRRTITGTTCN
 EETONLSRTIYLAEVOSKYKROIFESDYOSVDIYNKIDREL"
 BASE COUNT 255 a 135 c 164 g 222 t 1 others
 ORIGIN
 alignment_scores:
 Quality: 1300.00 Length: 240
 Ratio: 5.417 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.583
 alignment_block:
 US-09-528-682-1 x A04915 ..
 Align seg 1/1 to: A04915 from: 1 to: 777
 1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
 |||||
 55 AATGGCGACAGATTATACCGTGCCTAGACCCACAGATGAATAAA 104
 |||||
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||
 105 ACGTTCGAGGCTCTATGCCAGAGGCGCAATAGTACTTCGATGAG 154
 |||||
 34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||
 155 GAATCGAAATGAATATATCTTATGATCATCGGAGACAGACCAAC 204
 |||||
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||
 205 GCGTTTCTCAGATATGATGACGATATGTTTTCACCTCTCTAGTTGAG 254
 |||||
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 |||||
 255 AAGGCTCCTACTTACGACGACGCTATATATATCGGATATTCACCTACT 304
 |||||
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||
 305 ATATATATGTTATATGCGACACCAATATGTTTAAATGATGATGTA 354
 |||||
 101 LeuGlyValTyrSerProHisProTyrGluGlnIleValSerAlaLeuG 117
 |||||
 355 TTAGCGGTATACAGCCCTCACCATATGACAGAGAGGTTCTCGGTTAG 404
 |||||
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 |||||
 405 TGGAAATACATATTCATCATATATGATGATGATGATGATGATGATG 454

134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||
 455 TGATGTATGACAGATTATCATCGTATACAGGAGAAATATAGAGCCGGTATAC 504
 |||||
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
 |||||
 505 AGAAATCTGAATATAGCTCCGCGACGAGAGATGTTTACAGATTACAGAGTTT 554
 |||||
 167 eProProAspHisGlnAlaLeuPArgGluGluProTyrPheHisAlaP 184
 |||||
 555 CCACCGGATCACCACAGCTTGAGAGAGAAACCTGATTCATCATGACAC 604
 |||||
 184 roGlnGlyCysGlyAsnSerSerArgThrIlePheGlyAspThrCysAsn 200
 |||||
 605 CACAGGTTTGGAATTCATCATCAAGAACATCATCAGGTATATCTGTAT 654
 |||||
 655 GAGGAGACCCAGAAATCTGACACCAATATATCTCAGGGAATATCAATCAA 704
 |||||
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
 |||||
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 |||||
 705 AGTTAAGAGCGAGATATTTTCAGACTATCTGTCAGAGTTGACATATATA 754
 |||||
 234 snArgIleArgAspGluLeu 240
 |||||
 755 ACAGAAATCGGATGAAATTA 774
 seq_name: gb_pat:E00613
 seq_documentation_block:
 LOCUS E00613 777 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA encoding A component of heat labile toxin, LTA.
 ACCESSION E00613
 VERSION E00613.1 GI:2168892
 KEYWORDS JP 1986005097-A/1.
 SOURCE JP 1986005097-A/1.
 ORGANISM Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 777)
 REFERENCES
 1 Maikuru, B. H., Suteilbun, H. and Goodon, D. R.
 MICROBIAL METHOD
 Patent: JP 1986005097-A 1 10-JAN-1986;
 GIAXO GROUP LTD
 COMMENT
 OS Escherichia coli
 PN JP 1986005097-A/1
 PD 10-JAN-1986
 PE 12-DEC-1984 JP 1984262645
 PR 12-DEC-1983 GB 83 8333131
 PI MAIKURU BIKUTAA HEIZU, SUTELIBUN HAAFUODOO,
 PI GOODON UTRIMU ROSU
 PC CO7K15/04, A61K39/108, C12N1/20, C12N15/00//C12P21/00, (C12N1/20,
 PC C12R1:19),
 PC (C12N15/00, C12R1:19);
 CC strandedness: Single;
 CC topology: linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH key
 CDS 1..777
 /product="A component of heat labile toxin, LTA"
 FT mat_peptide 1..774
 /product="A component of heat labile toxin, LTA"
 FT mutation replace(236, 'T')
 /note="SDMI"
 FT Location/Qualifiers
 1..777
 /organism="Escherichia coli"

BASE COUNT 255 a 137 c 163 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1298.00 Length: 240
 Ratio: 5.408 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.167

alignment_block:
 US-09-528-682-1 x E00613 ..

Align seg 1/1 to: E00613 from: 1 to: 777

```

1 ASnglYAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
|||||
55 AATGGCGACAGATTATATACCGTCTGACTCTAGACCCCGAGATGAATAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACCTTCGGAGGCTTATGCCCCAGAGGCAATATAGTACTTCGATAGAG 154
34 lYthrglMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCAATGATATATATCTTATGATCACCAGAGAGAACACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||
205 GGCTTGTGAGATATGATGACGAGATATGTTCCACTTCTCTAGTTTGA 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||
255 AAGTGTCTACTTATGCGAGAGAGCTATATATATCAGATATATCCACTT 304
84 yrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTTATAGCGACAGACCAATATATGTTATGTTATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
|||||
355 TTAGGGGTATACAGCCCTCACCCTATATGAACAGAGTTTCTCGTTAG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
|||||
405 TGGAAATACCATATTTCTCAGATATATGATGATCGTTAATTTTGGTG 454
455 TGATGTATGACGATACATCGTAACAGGAATATAGAGCCGCTATTAC 504
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||
505 AGAAATCTCAATATATACCTCCGCGAGAGATGTTACAGATTACAGG 554
167 ePProAspHisGlnAlaThrArgGluIleProTyrPheHisAlaP 184
|||||
555 CCCACCGGATCGAAGCTTGGAGAGAGAACCTGATTCATCATGACAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAAAGGTTGTGAATTCATCAAGAACATACAGAGTGTACTGTAAAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
655 GAGGACACCCAGATCTGAGCACAAATATATCTCAGGGAATATCAATCA 704
217 sValTyrSerGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
705 AGTTAAGAGGAGATATTTTTCAGACTATACGTCAGAGGTTGACATATA 754
234 snArgIleArgAspGluLeu 240
|||||

```

755 ACAGAAATTCGAGATGAATTA 774

seq_name: gb_ba:EWDLTA

seq_documentation_block:

LOCUS EWDLTA 723 bp DNA linear BCT 23-APR-1996
 DEFINITION Plasmid EMD 299 (from E.coli) heat-labile enterotoxin A
 (LT-A) gene, 3' end.
 ACCESSION M57244 J05702
 VERSION M57244.1 GI:531190
 KEYWORDS enterotoxin; heat-labile enterotoxin.
 SOURCE Plasmid EMD 299 DNA.
 ORGANISM
 plasmids.

REFERENCE
 1 (bases 1 to 723)
 Tsuji,T., Inoue,T., Miyama,A., Okamoto,K., Honda,T. and Mawatani,T.
 A single amino acid substitution in the A subunit of Escherichia
 coli enterotoxin results in a loss of its toxic activity
 J. Biol. Chem. 265 (36), 22520-22525 (1990)
 91093102
 On Aug 20, 1994 this sequence version replaced gi:148531.
 COMMENT
 FEATURES
 source
 1..723
 /organism="Plasmid EMD 299"
 /plasmid="Plasmid EMD 299"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2464"
 1..723
 /gene="LT-A"
 <1..723
 /gene="LT-A"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AAB59161.1"
 /db_xref="GI:148532"
 /translation="MGDRLYRADSNPPDEIKRSGGLMRGHNEXEFGRTGNINLYDH
 ARGQTGVRYRDDGVSTSLRSNHLAGOSILSGSYTYIVYATAPMNFVNDVLG
 VYSPHPYQAKXSALAGIPYSQTYKRYRNFVIDERLGNREYRRTYRNINAPAE
 GYRLAGFPDHOAMREFPWIRHAPQGCNNSRTYTGDTCEETONLSTIYIREYSKY
 KRQIFSDYQSEVDIYNRIDEL"
 459

gene
 CDS

old_sequence

BASE COUNT 238 a 128 c 160 g 197 t
 ORIGIN

alignment_scores:
 Quality: 1293.00 Length: 240
 Ratio: 5.410 Gaps: 0
 Percent Similarity: 99.583 Percent Identity: 99.167

alignment_block:
 US-09-528-682-1 x EMDLTA ..

Align seg 1/1 to: EMDLTA from: 1 to: 723

```

1 ASnglYAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
|||||
1 AATGGCGACAGATTATATACCGTCTGACTCTAGACCCCGAGATGAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
51 ACCTTCGGAGGCTTATGCCCCAGAGGCAATATAGTACTTCGATAGAG 100
34 lYthrglMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
101 GAACCAATGATATATATCTTATGATCACCAGAGAGAACACAAACC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||

```



```

151 GGGCTTGTCAGATATGATGACGGATAGTTTCCACTCTCTGATTGAG 200
67 gserAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTGCCTCAGCTTAGCAGACAGCTATATATATCAGGATATCCACTACT 250
84 ylleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATATAGCCACACACCAAAATATGTTTAATGTTAATGATGA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
301 TTAGGCGGTATACAGCCCTCACCATATGACAGAGAGAGGTTCTCGGTTAG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTTPtyrArgValAsnPhcGly 134
351 TGGAAATACCATATTCCTCAGATATATGATGATGATACGTTAAATTTGGTG 400
134 aAlleAspGluArgLeuHisArgAsnArgS1uTyrArgAspArgTyrTyr 150
401 TGAATGATGAAACGATTTCATGTAACAGGAAATATAGACCGGATATTAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
451 AGAAATCTGAATATAGCTCCGCGCAGAGATGTTACAGATTAGACAGGTTT 500
167 eProProAspHisGlnAlaThrPargGluGluProTPrIleHisHisAlaP 184
501 CCCACCGGATCCACCAAGCTTGAGAGAAACCTGGATTCCATCATGTCAC 550
184 rGlnGlnCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CACAGAGTTGTGAAATTCATCAAGAAACATCACAGGTGATCTGTAAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
601 GAGGAGACCCAGAAATCGACACATATATCTCAGAGAAATATCAATCAA 650
217 sValIlyArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
651 AGTTAAAGGACAGATATTTTCAGACTATTCAGTCAGAGGTTGACATATATA 700
234 snArgIleArgAspGluLeu 240
701 ACAGAAATTCGGAGTGAATTA 720

seq_name: gb_pat:E03421

seq_documentation_block:
LOCUS E03421 1148 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding Lth of entero toxigenic Escherichia coli.
ACCESSION E03421.1 GI:2171637
VERSION E03421.1 JP 1992079898-A/1.
KEYWORDS JP 1992079898-A/1.
SOURCE Escherichia coli.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 1148)
AUTHORS Danbara,H. and Abe,A.
TITLE DNA AND RNA PROBE TO SIMULTANEOUSLY DETECT VIBRIO CHOLERAE AND
ENTEROTOXIGENIC ESCHERICHIA COLI AND DETECTION OF VIBRIO CHOLERA
AND ENTEROEXIGENIC ESCHERICHIA COLI USING THE SAME DNA AND RNA
PROBE
JOURNAL Patent: JP 1992079898-A 1 13-MAR-1992.
COMMENT
KITASATO INST:THE
OS Escherichia coli
PN JP 1992079898-A/1
PD 13-MAR-1992
PF 23-JUL-1990 JP 1990194208
PI DANBARA HIROFUMI, ABE AKIO
PC C1201/68,C07H21/02,C07H21/04,C12N15/11,C12N15/31,C1201/04, PC

```

```

C1201/10,
PC (C12N15/31,C12R1:19),(C1201/04,C12R1:63,C12R1:19): CC
strandedness: Double:
CC topology: linear:
CC hypothetical: No:
CC anti-sense: No:
CC *source: clone-pkAD008;
FH key Location/Qualifiers
FH misc-feature 1..1148
FT /note="DNA fragment encoding Lth of entero FT
FT sig_peptide 1..54
FT sig_peptide 55..774
FT mat_peptide 1..777
FT CDS /product="subunit A of Lth"
FT sig_peptide 774..837
FT mat_peptide 838..1145
FT CDS /product="subunit B of Lth"
FT CDS /product="precursor of subunit B of Lth"
FEATURES
source Location/Qualifiers
1..1148
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 399 a 205 c 229 g 315 t
ORIGIN
alignment_scores:
Quality: 1292.00 Length: 240
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.750
alignment_block:
US-09-528-682-1 x E03421 ..
Align seg 1/1 to: E03421 from: 1 to: 1148
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 AATGGCGACAAATTTATACCGTCTGACTCTACAGCCCGCAGATGAATAA 104
17 satGserGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 ACGTCCGGAGGCTTATGCCCCAGAGGCAATATGAGTCTTCGATGAG 154
34 lYthrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GAATCAATAGATATTAATCTTTATGATCAGCCGAGAGAACACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GCCTTTGTCAGATATGATGACGGATATGTTTCCACTCTCTTAGTTTGG 254
67 gserAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 AAGTGCCTCAGTACGACGACAGCTATATATATCAGATATTCACACTTACT 304
84 ylleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATATATATGTTATATAGCAGACACCAAAATATGTTTAATGTTATGATGA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTAGGCGGTATACAGCCCTCACCATATGAAACAGAGAGGTTCTCGGTTAG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTTPtyrArgValAsnPhcGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGGAAATACCATATTCCTCAGATATATGATGATGATACGTTAAATTTGGTG 454

```



```

/protein_id="BA25725.1"
/db_xref="GI:3062901"
/translation="MKNTTFEFTILASPIVANGDKYLRAPSDPDNFKPSGCTMPGQ

```

645 CCCACGGATCACCAAGCTTGAGAGAGAACCTGGATTTCATCATGAC 699

184 roGInGlycylAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||:::|||||
 695 CACAAGGTGTGGAATTCATCAAGAACATTTACAGATGATCTGTAA 744
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
 |||||:::|||||
 745 GAGGAGACCCAGCAATCTGAGCACAAATATATCTCAGCAAAATATCAATCAAA 794
 217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||:::|||||
 795 AGTTAAGAGCGACATATTTTCAGACTATCATGTCAGAGGTGACATATATA 844
 234 snArgIleArgAspGluLeu 240
 |||||:::|||||
 845 ACAGAAATTCGGGATGAATTA 864

seq_name: gb_ba:ECOTOXA

seq_documentation_block:

LOCUS ECOTOXA 934 bp DNA linear BCT 26-APR-1993
 DEFINITION E.coli toxA gene encoding subunit A of heat-labile enterotoxin.
 ACCESSION K01995
 VERSION K01995.1 GI:148027
 KEYWORDS enterotoxin; heat-labile enterotoxin; toxA gene.
 SOURCE E.coli H10407 (serotype 078:H11) plasmid DNA, clone pJY27.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE
 1 (bases 1 to 934)
 AUTHORS Yamamoto,T., Tamura,T. and Yokota,T.
 TITLE Primary structure of heat-labile enterotoxin produced by
 Escherichia coli pathogenic for humans
 JOURNAL J. Biol. Chem. 259, 5037-5044 (1984)
 MEDLINE 84185610

COMMENT This heat-labile enterotoxin is pathogenic for humans (Lth).
 Compared in [1] with Ltp A and Ct A (Ltp-pathogenic for piglets,
 Ct-choleera toxin).

FEATURES

source

1..934

/organism="Escherichia coli"

/db_xref="taxon:562"

158..934

/note="heat-labile enterotoxin A prepeptide"

/codon_start=1

/transl_table=1

/protein_id="AAA24685.1"

/db_xref="GI:148028"

/translation="MKNTFFFFILLASPLYANGDKLYRADSRPPDEIKRGGIMPRG
 HNEYFDKGTOMNINLYDHARGTQGFVARYDGYVSTLSLSAHLAGOSILSGSTYY
 IYVATAPNMNVNDVLGVSPHPYEDQVSLAGIPLYSQIIGWYRVNFGVIDERLHRN
 REYDRYRNINIAEDGYRLAGFPDPHQAAREPMTLHHAPCGGDSRRITGDTGN
 EETONLSITVYLRKQSKYKRQIFSDYQSEVDIVNRREL"

158..211

/note="heat-labile enterotoxin A signal peptide"

212..787

/product="heat-labile enterotoxin A1"

794..931

/product="heat-labile enterotoxin A2"

BASE COUNT 298 a 165 c 186 g 285 t

ORIGIN 46 bp upstream of PstI recognition pattern.

alignment_scores:
 Quality: 1287.00 Length: 240
 Ratio: 5.362 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.333

alignment_block:
 US-09-528-682-1 x ECOTOXA ..

Align seg 1/1 to: ECOTOXA from: 1 to: 934

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleLeu 17
 |||||:::|||||
 212 AATGGCGACAAATTTATACCGTGCTGACTAGACCCCGAGATGAATAA 261
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||:::|||||
 262 ACGTTCCGAGGCTTATGCCAGAGGCGCATATATGATGCTTCGATAGAG 311
 34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||:::|||||
 312 GAACTCAATATGAATTTAATCTTTATGATCAGCCGAGAGAACACCAAC 361
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||:::|||||
 362 GGCTTGTGCAGATATGATGACGATATGTTCCACTTCCTTAGTTGAG 411
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 |||||:::|||||
 412 AAGTGCTCACTTAGCAGACAGCTATATATATCAGATATTCACCTTACT 461
 84 YrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||:::|||||
 462 ATATATATGTTATAGCGACAGACCAAAATATGTTTATATTAATGATGA 511
 101 LeuGlyValTyrSerProHisProTyrGlnGlnGluValSerAlaLeuG 117
 |||||:::|||||
 512 TTAGCGGTATACAGCCCTCACCATATGAACAGAGGTTTCGCGTTAG 561
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
 |||||:::|||||
 562 TGGAAATACCATATATTCAGATATATGATGATGATCTGTTAATTTGGTG 611
 134 alleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||:::|||||
 612 TCAATTATGACAGCATTTACATCGTAACAGGAAATATGACACCGGTATAC 661
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 |||||:::|||||
 662 AGAAATCTGAATATATGCTCCGGCAGAGATGTTACAGATTAGCAGATWT 711
 167 eProProAspHisGlnAlaThrArgGlnGluProTyrPheHisAlaP 184
 |||||:::|||||
 712 CCCACCGGATCACCAAGCTTGAGAGAGAACCTTGATTCATCATGACAC 761
 184 roGInGlycylAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||:::|||||
 762 CACAAGGTGTGGAATTCATCAAGAACATTTACAGTATCTGTAA 811
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
 |||||:::|||||
 812 GAGGAGACCCAGCAATCTGAGCACAAATATATCTCAGCAAAATATCAATCAAA 861
 217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||:::|||||
 862 AGTTAAGAGCGACATATTTTCAGACTATCATGTCAGAGGTGACATATATA 911
 234 snArgIleArgAspGluLeu 240
 |||||:::|||||
 912 ACAGAAATTCGGGATGAATTA 931

seq_name: gb_ba:S60731

seq_documentation_block:

LOCUS S60731 1275 bp DNA linear BCT 29-JAN-2002
 DEFINITION heat-labile enterotoxin A subunit, heat-labile enterotoxin B
 subunit [Escherichia coli, 21d, Genomic, 1275 nt].
 ACCESSION S60731
 VERSION S60731.1 GI:408994
 KEYWORDS
 SOURCE Escherichia coli 21d.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1 to 1275)
 AUTHORS Inoue,T., Tsuji,T., Koto,M., Imanura,S. and Miyama,A.
 TITLE Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic *Escherichia coli* is identical to that of human strain H 10407
 JOURNAL FEMS Microbiol. Lett. 108 (2), 157-161 (1993)
 MEDLINE 93352225
 PUBMED 8486242
 REMARK Genbank staff at the National Library of Medicine created this entry [NCBI glibseq 131491] from the original journal article. This sequence comes from Fig. 2.
 FEATURES
 source Location/Qualifiers
 1..1275
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 68..844
 CDS
 /note="This sequence comes from Fig. 2: conceptual translation presented here differs from translation in publication: Ltc A subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin A subunit"
 /protein_id="AAC60440.2"
 /db_xref="GI:18390349"
 /translation="MKNTFFFIILASPLVANGDKLYRADSRPDEIKRSGLMPRG HNEYFDRTQMINILYDHAQGTGVRDQGVSTSLRSALHAGOSILSGYSTY IYVATPMFNVNDVGLGVSPHYEVSALGIPYSQIYGVNVRNGVIDERLARN REYRDYRNINLIPAEQGRVLRAGPPRPHQARREPMTHIAQGGDSRTITGDTCN EETQNLSTIYLKRYOSKVRKQIFSDYSQSEVDIYNIRNEL"
 841..1215
 CDS
 /note="This sequence comes from Fig. 2: Ltc B subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin B subunit"
 /protein_id="AAC60441.1"
 /db_xref="GI:408996"
 /translation="MNKVCYVLPFTALSLCAGAPOSTELCSEYRNTQIYTIINDK ILSYEMNAGKRENVITFEKSGATFOVEVPSQIHDSQKAIERMKDRLRIYLMETK IDLCVNMNKTPTNSIAIAISMEN"
 BASE COUNT 436 a 227 c 248 g 364 t
 ORIGIN
 alignment_scores:
 Quality: 1287.00 Length: 240
 Ratio: 5.362 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.333
 alignment_block:
 US-09-528-682-1 x S60731 ..
 Align seg 1/1 to: S60731 from: 1 to: 1275
 1 ASnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
 |||||:|||||:
 122 AATGGCACAATAATATACCGTGTGACTGTAGACCCACAGATGAATAAA 171
 |||||:
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyIuTYrPhAspArg 34
 |||||:
 172 ACGTTCGGAGGCTTATGCGCACAGGGCATTAAGTACTTCCATGAG 221
 |||||:
 34 LYrHGlMetAsnIleAsnLeuTYrAspHisAlaArgGlyThGlnThr 50
 |||||:
 222 GAACCTCAATGAATTAATCTTATGATCAGCGAGAGAACCAACC 271
 |||||:
 51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuAr 67
 |||||:
 272 GCGTTTGTCAGATATGATGACGATATGTTCCACTTCTTACTGTTGAG 321
 |||||:
 67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTYrSerThrTYrT 84
 |||||:
 322 AAGTGTCACTTACGACGACAGCTATATATATCAGGATATTTCCACTTACT 371

84 yIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||:
 372 ATATATATGTTATGACGACGACCAATAATGTTATGTTAATGATGATA 421
 |||||:
 101 LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeuG 117
 |||||:
 422 TTAGCGGTATACAGCCCTCACCATATGAAACAGAGCGTTCCTCGCTTAC 471
 |||||:
 117 yGlyIleProTYrSerGlnIleTYrGlyTYrPTyrArgValAsnGly 134
 |||||:
 472 TGGATACCATATTCATCATATATGATGATGATCGTTATATTTGCTG 521
 |||||:
 134 alIleAspGluArgLeuHisArgAsnArgGlyTYrArgAspArgTYrTYr 150
 |||||:
 522 TGATGTATGAACGATTTACATCGTAAACGAGCAATATAGACCGGTATAC 571
 |||||:
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
 |||||:
 572 AGAATCTGAATATAGCTCCGCGAGAGATGCTTACAGATTAACGAGT 621
 |||||:
 167 eProProAspHisGlnAlaTrpArgGluGluProTYrIleHisAlaP 184
 |||||:
 622 CCCACCGGATCCACCACTTGAGAGAAACCCTGGATTCATCATATCAC 671
 |||||:
 184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||:
 672 CACAAGGTGTGGAGATTTCATCAAGAACATTAACAGGTGATCTTGNAT 721
 |||||:
 201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSer 217
 |||||:
 722 GAGGAGACCCGATCTGACGACAAATATCTCAGAAATATCAATCAAA 771
 |||||:
 217 sValIleArgGlnIlePheSerArgTYrGlnSerGluValAspIleTYr 234
 |||||:
 772 AGTTAAGAGCGATATTTTCAGACTATCATGCTCAGAGGTGACATATATA 821
 |||||:
 234 snArgIleArgAspGluLeu 240
 |||||:
 822 ACAGAAATTCGAATGAATGTA 841
 |||||:
 seq_name: gb_pat: A16419
 seq_documentation_block:
 LOCUS A16419 711 bp DNA linear PAT 06-JUL-1995
 DEFINITION A subunit of a labile toxin LT1.
 ACCESSION A16419
 VERSION A16419.1 GI:641008
 KEYWORDS
 Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 ORGANISM
 1 (bases 1 to 711)
 REFERENCE
 1 IMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN
 LT1, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES
 Patent: WO 93/3202-A 42 08-JUL-1993;
 JOURNAL
 TITLE
 LOCATION/Qualifiers
 1..711
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..711
 /gene="LT1 subunit A"
 <1..711
 /transl_table=11
 /codon_start=1
 /protein_id="CAA01280.1"
 /db_xref="GI:641009"
 /translation="NGDRLYRADSRPDEIKRSLMPRGNEYFDRTQMINILYDHA RCTQGFVARDYGVSTSLRSALHAGOSILSGSLTYIYIAMPFNVNDIVSYSP HPEQEVNLAGIIPYSQIYGVNVRNGVIDERLARNREYRDYRNINLIPAEQGRV LRAGPPRPHQARREPMTHIAQGGDSRTITGDTCN EETQNLSTIYLKRYOSKVRKQIFSDYSQSEVDIYNIRNEL"
 CDS
 1..711

BASE COUNT 234 a 123 c 156 g 198 t
 ORIGIN

alignment_scores:

Quality: 1183.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:

US-09-528-682-1 x A16419

Align seg 1/1 to: A16419 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
|||||
1 AATGGCGACAGATTATACCGTCTGACTAGACCCCGAGATGAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
51 ACCTTTCCGAGCTTATGCCAGAGGT...AATGAGTACTTGATAGAG 97
34 LyrHrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
98 GAACCTCAATGATATTAATCTTATGATCAGCGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
148 GGCTTTGTCAGATATGATGACGGATATGTTCCACTTCTCTGTTGAG 197
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
198 AATGCTCTACTTAGCAGACAGATATATATATACGATATCTACTACT 247
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
248 TATATATCGTTATAGCA.....AATATGTTTATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
289 AATAGGCTATACAGCCCTCACCATTATGAACAGAGGTTTCTCGTTAGG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhGly 134
|||||
339 TGGAAATACCATATTCTCAGATATATGATGATCGGTAAATTTTGCTG 388
134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
389 TGAATGATGACGATTACATCATACAGGAATATAGAGACCGGTATTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
439 AGAAATCTGATATAGTCCGCGAGAGGATGTTTACAGATTACAGAGTTT 488
167 eProProAspHisGlnAlaThrPargGluGluProTyrPheHisAlaP 184
|||||
489 CCACCGGATACCAAGCTTGAGAGAGAACCCCTGATTCATCATACGAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
539 CACAAGGTTGTGAGATTCTCATCAAGAACATACAGGTGATCTGTAAT 588
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
589 GAGGAGACCCAGATCTGAGCACAAATATATCTCAGGAAATATCAATCAAA 638
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
639 AGTTAAGAGGCAATATTTTTCACACTATCATGTCAGAGGTTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
|||||

```

689 ACAGAAATCGGATGATTA 708

seq_name: gb_pat:AR118597

seq_documentation_block:

LOCUS AR118597 711 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 5 from patent US 6149919.
 ACCESSION AR118597
 VERSION AR118597.1 GI:14100507
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 711)

AUTHORS Domenighini,M., Rapunoli,R., Pizza,M. and Hol,W.
 TITLE Immunogenic detoxified mutants of cholera toxin and of the toxin
 LT, their preparation and their use for the preparation of vaccines
 JOURNAL Patent: US 6149919-A 5 21-NOV-2000;
 FEATURES Location/Qualifiers
 source 1..711

BASE COUNT 234 a 123 c 156 g 198 t
 ORIGIN

alignment_scores:

Quality: 1183.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:

US-09-528-682-1 x AR118597

Align seg 1/1 to: AR118597 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
|||||
1 AATGGCGACAGATTATACCGTCTGACTAGACCCCGAGATGAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
51 ACCTTTCCGAGCTTATGCCAGAGGT...AATGAGTACTTGATAGAG 97
34 LyrHrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
98 GAACCTCAATGATATTAATCTTATGATCAGCGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
148 GGCTTTGTCAGATATGATGACGGATATGTTCCACTTCTCTGTTGAG 197
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
198 AATGCTCTACTTAGCAGACAGATATATATATACGATATCTACTACT 247
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
248 TATATATCGTTATAGCA.....AATATGTTTATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
289 AATAGGCTATACAGCCCTCACCATTATGAACAGAGGTTTCTCGTTAGG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhGly 134
|||||
339 TGGAAATACCATATTCTCAGATATATGATGATCGGTAAATTTTGCTG 388
134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
389 TGAATGATGACGATTACATCATACAGGAATATAGAGACCGGTATTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||

```


439 AGAATGTGATATAGCTCCGGCAGAGATGTTACAGATTACGAGTTT 488
 167 eProPAsphigInlaIaTrPARgLUgLUProTrIleHsIaIaP 184
 |||||||
 489 CCCACCCGATCACCAGCTTGAGAGAAACCTGGATTCATCATGCAC 538
 184 roGInglYcysGlyAsnSerSerArgThrIleHrgIAspThrCysAsn 200
 |||||||
 539 CACAAAGGTGTGAGATTTCATCAAGAACATCAGAGTGATCTTGAAT 588
 201 GluLUhrgInaSnLeuSerThrIleTyrlEuargLUgLUgInserLy 217
 |||||||
 589 GAGAGACCCAGATCTGAGACAAATATCTCAGGAAATTCATCAACAA 638
 217 sVallyArGInIlePheSerAspTyrgInSerGluValAspIleTyra 234
 |||||||
 639 AGTTAAGGACAGATATTTTCAGACTATCAGTCAAGAGTGCATATATA 688
 234 snArGIIeaTgAspGluLeu 240
 |||||||
 689 ACAGAAATTCGGGATGAATTA 708

seq_name: gb_ba:ECCELTA

seq_documentation_block:

LOCUS ECCELTA 868 bp DNA linear BCT 12-SEP-1993
 DEFINITION E. coli gene elta encoding the A subunit of the heat-labile
 enterotoxin (LT).
 ACCESSION V00275.1 GI:41339
 VERSION V00275.1 GI:41339
 KEYWORDS enterotoxin; signal peptide.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1 to 868)
 AUTHORS Slicer,E.K. and Noble,J.A.
 TITLE Escherichia coli heat-labile enterotoxin. Nucleotide sequence of
 the A subunit gene
 JOURNAL J. Biol. Chem. 257 (10), 5716-5721 (1982)
 MEDLINE 82167425

FEATURES
 source Location/Qualifiers

1..868
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 56..>868
 /product="messenger RNA"
 90..143
 90..854
 /codon_start=1
 /transl_table=11
 /product="enterotoxin subunit A"
 /protein_id="CAA23532.1"
 /db_xref="GI:41340"
 /translation="MKNTFFIFILLASNGRLYPADRPDEIKRSLMPRG
 NEYFDRTQMINILYDHAAGTGTGFRYDDGVYSLSIRSAHLAGVYLSGLTYI
 IVIANPENVNDIVSYSPRHYRQEVSAIGIYISQITGRYRNFGVIDERLNRNEYR
 DRYRNLIAPREDGIRLAGFPDHDAMREEPHIAHPQGGDSSRTTGTGDCNETQ
 NLSTYILEYKSVKROJFSDIQSEVDIYNRIIRDEL"

BASE COUNT 284 a 142 c 179 g 263 t
 ORIGIN

alignment_scores:
 Quality: 1183.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:
 US-09-528-682-1 x ECCELTA ..

Align seg 1/1 to: ECCELTA from: 1 to: 868

1 AsnGlyAspArgLeuTyraAlaAspSerArgProPAspGluIleLy 17
 |||||||
 144 AATGCCACAGATTAATACCGTGTGACTTGACCCACAGATGAATATAA 193
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrrPheAspArg 34
 |||||||
 194 ACGTTCCGAGTCTTATGCCAGAGT...AATGAGTACTTCGATAGAG 240
 34 LyThrGlnMetAsnIleAsnLeuTyraSphIsaIaArgGlyThrGlnThr 50
 |||||||
 241 GAATCAATATGAATTAATCTTATGATCAACGGAGAGAACACAAACC 290
 51 GlyPheValArgTyraAspAspGlyTyraValSerThrSerLeuSerLeu 67
 |||||||
 291 GCGTTTGCAGATATGATGACGATATGTTTCCACTTCTTACGTTGAG 340
 67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrrSerThrTyrr 84
 |||||||
 341 AAGTGCTCACTTACAGAGAGTATATATTCAGAGATATTCACCTACTA 390
 84 yTleTyrrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
 |||||||
 391 TATATATCGTTATAGCA.....AATATGTTTAATGTTAATGATGTA 431
 101 LeuGlyValTyrrSerProHisProTyrgLUgLUgInIuValSerAlaLeuG 117
 |||||||
 432 ATTAGCGATACAGCCCTCCACCATGACAGAGAGGTTTGTGGCTTAGG 481
 117 yGlyIleProTyrrSerGlnIleTyrgLUgLUgTyrrTyraValAsnPheGly 134
 |||||||
 482 TGAATACCATATTCCTCAGATATATGATGTGATACGTTATATTTGGTG 531
 134 alIleAspGluArgLeuHisArgAsnArgGluTyrrTyraAspArgTyrr 150
 |||||||
 532 TGATTTGATGAACGATTCATGCTGAACAGGGAATAGACCGGTATATAC 581
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrrTyraValaIaGly 167
 |||||||
 582 AGAATTCGAATATAGCTCCGGCAGAGAGATGTTACAGATTAGCAAGTTT 631
 167 eProPAsphigInlaIaTrPARgLUgLUProTrIleHsIaIaP 184
 |||||||
 632 CCCACCCGATCACCAGCTTGAGAGAAACCTGGATTCATCATGCAC 681
 184 roGInglYcysGlyAsnSerSerArgThrIleHrgIAspThrCysAsn 200
 |||||||
 682 CACAAGGTGTGAGATTTCATCAAGAACATCAGAGTGATCTTGAAT 731
 201 GluLUhrgInaSnLeuSerThrIleTyrlEuargLUgLUgInserLy 217
 |||||||
 732 GAGAGACCCAGATCTGAGACAAATATCTCAGGAAATTCATCAACAA 781
 217 sVallyArGInIlePheSerAspTyrgInSerGluValAspIleTyra 234
 |||||||
 782 AGTTAAGGACAGATATTTTCAGACTATCAGTCAAGAGTTGCATATATA 831

seq_name: gb_ba:AF452584

seq_documentation_block:

LOCUS AF452584 777 bp DNA linear BCT 14-JAN-2002
 DEFINITION Vibrio cholerae strain 1322-69 CtxA (ctxA) gene, complete cds.
 ACCESSION AF452584
 VERSION AF452584.1 GI:18141148
 KEYWORDS
 SOURCE Vibrio cholerae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 REFERENCE 1 (bases 1 to 777)

AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1..777
/organism="Vibrio cholerae"
/strain="1322-69"
/db_xref="taxon:666"
/note="Identical sequence found in Vibrio cholerae strains
571-88 (serogroup O105) and 203-93 (serogroup O141)
serogroup: O37"
1..777
/gene="ctxA"
1..777
/gene="ctxA"
/codon_start=1
/transl_table=1
/product="ctxA"
/protein_id="AA060525.1"
/db_xref="GI:18141149"
/translation="MWKIIFFVFIPLSFSFYANDKLYRADSRPDEIKOSGGLMPRG
QNEYDRGTOMNINILYDHARGTQTFVHDDGYSTSLSAHLVGGTILSGHSTY
IYVITADNMENVDNLGAYSPHDEQVSAAGIPIYQIYGVYHVFHGVDEQLHRN
RGYRRYVSNLDIAPADYGLAGFPPEHRAAREPWHHAAPGCGNAPRSMSTCD
EKTQSLGYKFLDEYQSKYKROIFSGYQSDIDITHNIRKDEL"
BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN
alignment_scores:
Quality: 1093.00 Length: 240
Ratio: 4.836 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 82.083
alignment_block:
US-09-528-682-1 x AF452584 ..
Align seg 1/1 to: AF452584 from: 1 to: 777
1 AsnglyAspargleuTYrArgAlaAspSerArgProProAspGluIlely 17
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGATATAGTTATATCGGCAGATTCTAGACCTCGATGATAATAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTYrPheAspArg 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
105 GCAGTCAGTGGTCTTATGCCAAGACAGCAAGTGAAGTACTTGCACGAG 154
34 lythrMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GACTCTCAATGAATATCACTTATGATCATGCAGAGGAACTCAACAGC 204
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
205 GGATTTGTTAGGCAGATGATGATAGTTTCCACCTCATTTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTYrSerThrTYr 84
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
255 AAGTGGCCACTAGTGGTCAACATAATGTCGTGCATCTACTATAT 304
84 yrlleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
305 ATATATATGTTATAGCCACTGACCCCAACATGTTTAACTTAATGATGA 354

101 LeuGlyValTYrSerProHisProTYrGluGluGluValSerAlaLeuG 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
355 TTAGGGGCATACAGTCTCTCATCCAGATGACAGACAGAGTTCTGCTTAG 404
117 yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhGly 134
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
405 TGGGATTTCCCTACTCCCAATATATGATGATGATTCAGTTCATTTGGG 454
134 allLeaspGluArgLeuHisArgAsnArgGluTYrTYrArgAspArgTYr 150
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
455 TGCTTGATGACATTTATCATGTAATATGAGGCGTACAGAGATATATAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
505 AGTAACTTATATATGCTCCAGCAGCAGATGATGATGATGAGTTCAGGTT 554
167 eProProAspHisGlnAlaTrpArgGluGluProTYrIleHisAlaP 184
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
555 CCTCCGAGCATTAGAGCTTGAGGAGAGCCGTGATTCATCATGACAC 604
184 rGluGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
605 CGCGGGTGTGGAAATGCTCAAGATCATGATGATGATGATGATGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerly 217
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
655 GAAACCAACCAAGTCTAGTGTAAATTCCTTGAGCAATACCAATCTAA 704
217 sValIleArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
705 AGTTAAAGCAATATTTTCAGGCTATCATCTGATTTGATATACATA 754
234 snArgIleArgAspGluLeu 240
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
755 ATAGAAATTAGATGATTA 774
seq_name: gb_ba:AF463400
seq_documentation_block:
LOCUS AF463400 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 571-88 CtxA (ctxA) gene, complete cds.
ACCESSION AF463400
VERSION AF463400.1 GI:18448888
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1..777
/organism="Vibrio cholerae"
/strain="571-88"
/db_xref="taxon:666"
/note="serogroup: O105"
1..777
/gene="ctxA"
1..777
/gene="ctxA"


```

/codon_start=1
/translation=1
/product="ctxa"
/protein_id="AA169944.1"
/db_xref="GI:18448891"
/translation="MWKILFEFFIELSFVSANDKLYADSRPPEIKOSGGLMPRG
QHEYEDRGOMINILYDHAIGTGTGVRHDDGVYSISLSRSHLVGOTILSGHSTYY
IYVATAPMNFVNVDLGAISPDEOEVSAIGIPIYSQIYGMVRYHFGVLDLQHRN
RGYRDYVSNLDIAPADGVGLAGPPEHRAWREPEWHHAPPGCCNAPRSSMSNTCD
EKTOSIGVAFLEDEYOSKVKROJFSCGOSIDIDHNRKIDEL"
BASE COUNT      243 a      129 c      169 g      236 t
ORIGIN

alignment_scores:
  Quality: 1093.00      Length: 240
  Ratio: 4.836          Gaps: 0
  Percent Similarity: 94.167      Percent Identity: 82.083

Alignment_block:
US-09-528-682-1 x AF463400

Align seg 1/1 to: AF463400 from: 1 to: 777

1  AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
55 AATGATGATAGTATATACGGCAGATTCTAGACCTCTGATGAATAATAA 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17 SARGSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 GCATCTAGGTGTCTTATGCCAAGAGCAGAAATGACTTGCACGAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 GTACTCAATGATAATCAACCTTATGTCATGCAAGGAGACCTACAGC 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 GGATTTGTAGGCACGATGATGATGTTCCACCTCAATAGTTAGTTGAG 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255 AAGGCCCACTTAGTGGCTCAAACTATGCTGCTCATTTACTATTA 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspAla 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 ATAATATATGTTATAGCCACCTGACCCACACATGTTAAAGTAAATGATGTA 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 LeuGlyValTyrSerProHisProTyrGlnGlnIleValSerAlaLeuGlu 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 TTAGGGGATACAGTCCCTCATCCAGATGAACAAGATTTCTGCTTAGG 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyY 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 TGGGATTCACATACCTCCAAATATATGATGCTATCGATCATTTGGGG 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 aliLeaAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 TGCCTTGATGAACAATTCATCGTAAATAGGGCTACAGAGATGATGATTAAC 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 AGTAACCTTAGATATTGCTCCAGCAGAGATGTTATGATGGCAGGTTT 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 eProProAspHisGlnAlaTyrPargGluGluProTyrIleHisAlaP 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
555 CCCGCCGAGCATAGAGCTTGAGGGAAGACCCGCGATTCATCATGACAC 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 roGlnGlyCysGlyAsnSerSerArgTyrIleThrGlyAspThrCysAsn 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
605 CGCCGGGTGTGGGAATGCTCCAAAGATCATGATGATTAATCTTGGCAT 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerI 217

```

```

|||||.....:|||||
655 GAAAAAACCAAGCTAGGTGTAATAATTCCTTGACGAATACCAATCTAA 704
   |||.....:|||||
217 svallYsrGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
   |||.....:|||||
705 AGTAAAGACAAATATTTTCAGGCTATCAATCTGATATGATACACATA 754
   |||.....:|||||
234 snArgIleArgAspGluLeu 240
   |||.....:|||||
755 ATAGATTAAGCATGAATTA 774

seq_name: gb_ba:AF463401

seq documentation block:
LOCUS      AF463401              777 bp    DNA    linear    BCT 31-JUN-2002
DEFINITION Vibrio cholerae strain 203-93 CtxA (ctxa) gene, complete cds.
ACCESSION  AF463401
VERSION    AF463401.1 GI:18448890
KEYWORDS
SOURCE
ORGANISM   Vibrio cholerae.
            Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE  1 (bases 1 to 777)
            Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
            Sulakvelidze,A. and Sozhamannan,S.
            Genetic Analysis of the Virulence Regions, ctxA f prophage and
            Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
            Serogroup Strains of Vibrio cholerae
            Unpublished
            2 (bases 1 to 777)
            Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
            Sulakvelidze,A. and Sozhamannan,S.
            Direct Submission
            Submitted (27-DEC-2001) Department of Epidemiology and Preventive
            Medicine, University of Maryland School of Medicine, and VA
            Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
            Camden Street, Suite 675, Baltimore, MD 21201, USA
            Location/Qualifiers
            1..777
               /organism="Vibrio cholerae"
               /strain="203-93"
               /db_xref="taxon:666"
               /note="serogroup: O141"
            1..777
               /gene="ctxa"
               /gene="ctxA"
               /gene="ctxa"
               /codon_start=1
               /transl_table=11
               /product="ctxa"
               /protein_id="AA169944.1"
               /db_xref="GI:18448891"
               /translation="MWKILFEFFIELSFVSANDKLYADSRPPEIKOSGGLMPRG
               QHEYEDRGOMINILYDHAIGTGTGVRHDDGVYSISLSRSHLVGOTILSGHSTYY
               IYVATAPMNFVNVDLGAISPDEOEVSAIGIPIYSQIYGMVRYHFGVLDLQHRN
               RGYRDYVSNLDIAPADGVGLAGPPEHRAWREPEWHHAPPGCCNAPRSSMSNTCD
               EKTOSIGVAFLEDEYOSKVKROJFSCGOSIDIDHNRKIDEL"
BASE COUNT      243 a      129 c      169 g      236 t
ORIGIN

alignment_scores:
  Quality: 1093.00      Length: 240
  Ratio: 4.836          Gaps: 0
  Percent Similarity: 94.167      Percent Identity: 82.083

Alignment_block:
US-09-528-682-1 x AF463401

Align seg 1/1 to: AF463401 from: 1 to: 777

1  AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


55 AATGATGATAGTATATTCGGGCGAATTCAGACCTCCGTGATGAATAA 104
17 sargserglyyleumetproarglyhisnsglytyrphesparpg 34
105 GCAGTCAGGTGCTCTTATGCCAAGACAGACAGATGACTTATGACCGAG 154
34 lythrlnmetasnleasnleutyrsphisalaargglythrlnthr 50
155 GTACTCAATGATATATCACTTATGATCATGCAAGAGAACACAGACG 204
51 glyphevalargtyrpspsglytyrvalserthrserleuserleuar 67
205 GGATTTGTAGGACGATGATGATGTTCCACCTCAATTAATTTGAG 254
67 gseralhisleualaglylinserrileuserglytyrserthrtyr 84
255 AAGTCCCACTTAGTGGGTCAACTATATTTCTGGTCATTTCTTATTT 304
84 yrltyrvalillealathralaprosnmetpheasnvalasnspval 100
305 ATATATATGTTATAGCCACCTGACCAACATGTTTAACGTTAATGATGTA 354
101 leuglyvaltyrserprohisprotyrlyuglnlvalseralaleugl 117
355 TTAGGGGATACAGTCTCATCCAGATGACACAGAGTTTCTGCTTAGG 404
117 yglylleptryserglniletyrglytyrtyrargvalasnphelyv 134
405 TGGGATTCATACCTCCAAATATATGATGATGATGATGATTTTGGGG 454
134 alileasppluargleuhsargasnargglytyrargpsargtyrtyr 150
455 TGCTTGATGACAAATTCATCATAGGCGCTACAGAGATGATATATAC 504
151 argsnleuasnllealaproalagluaspglytyrargleualaglyph 167
505 AGTACCTTAGATGATTCCTCCAGACAGATGATGATGATGATGATGTT 554
167 eproprophisglnalatrparglyuglnprotyrphlehisalalap 184
555 CCCGCCGAGCATAGCTTGAGGAGAACCCGCTGATTCATCATGAC 604
184 roglnglycysglyasnserargthrilerhglyaspthrcysasn 200
605 CGCGGGGTTGTGGGATGCTCCAAAGATCATGATGATGATGATGATGAT 654
201 glugluthrglnasnleuserthrilerhlyleuarargglytyrlnserly 217
655 GAAAAAACCCAAAGCTAGGTGTAATAATCTCTGACGATACCAATCTTA 704
217 svallysrarglnlepheserargtyrlnserglyvalaspilleytyr 234
705 AGTTAAAGACAAATATTTTCAGGCTATCATGATGATGATGATGATGAT 754
234 snarglleargpsgluileu 240
755 ATAGATTAAGATGAATTA 774

seq_name: gb_ba:VIBCTXABA

seq_documentation_block:

LOCUS VIBCTXABA 1372 bp DNA linear BCT 01-FEB-2000

DEFINITION Vibrio cholerae ctxA gene and ctxB gene for cholera toxins, complete cds.

ACCESSION D30052.1 GI:487332

VERSION 3.0

KEYWORDS ADP-ribosyltransferase; cholera toxin.

SOURCE Vibrio cholerae (sub-species O37, strain S7) (library: plasmid) DNA, clone pKX10.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio. 1 (bases 1 to 1372)

REFERENCE 1 Yamamoto, K., Do, V.G., Xu, M., Iida, T., Miwatani, T., Albert, M.J. and

TITLE Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio cholerae strains 854 (O139-bengal) and S7 (O37) from two outbreaks unpublished (1994)

REFERENCE 2 (bases 1 to 1372)

AUTHORS Yamamoto, K.

COMMENT Direct Submission Submitted (29-Apr-1994) Koichiro Yamamoto, Osaka University, Institute for Microbial Diseases, Department of Bacterial Infections: 3-1 Yamadaoka, Suita, Osaka 565, Japan (Tel:06-879-4066) Submitted (29-Apr-1994) to DDBJ by: Koichiro Yamamoto Institute for Microbial Diseases Osaka University 3-1 Yamadaoka Suita, Osaka 565 Japan

FEATURES source 1. 1372

gene /organism="Vibrio cholerae" /strain="S7" /sub_species="O37" /db_xref="taxon:666" /clone_lib="plasmid" 113..889 /gene="cxtA" 113..889 /gene="cxtA" /codon_start=1 /product="cholera toxin" /protein_id="BAA06288.1" /db_xref="GI:808899" /translation="WVKILFVFIFLSEFSYANDDKLYRADSRPDEIKOSGLMPRG ONEYFDRGTOMINILYDHARGTGTGFEVHDGQVYSISLSRAHLVGGTILSGSTYY IYVITATNMEVNVNVLGASPHPEDEVSALGIPYSGIYGWYRVHGFVDEQLHRN RGYRDYRVSNLDIAPADGAGGPRPRHMRPREPWTHNAPPCGNAPRSMNCTCD EKTQSLGVKFLDEYGVSKRQIFSGYSDIDTHRIKDEL" 113..166 /gene="cxtA" 886..1260 /gene="cxtB" 886..1260 /gene="cxtB" /codon_start=1 /transl_table=1 /product="cholera toxin" /protein_id="BAA06289.1" /db_xref="GI:808900" /translation="MIKILFGVFYVLLSAYVHGTQNTIDCAEYHNTQIHLNDK ILSTESLAGNREMAIITFKNGATQVYPPSQHDSKKAIEKKKDLRIAYILEAK VEKLCVWNNKTPHAIATISMAN" 886..948 /gene="cxtB" 886..948

CDS sig_peptide 113..166

gene 886..1260

CDS 886..1260

sig_peptide 113..166

BASE COUNT 449 a 217 c 269 g 437 t

ORIGIN

alignment_scores: Quality: 1093.00 Length: 240 Ratio: 4.836 Gaps: 0 Percent Similarity: 94.167 Percent Identity: 82.083

alignment_block: US-09-528-682-1 x VIBCTXABA ..

Align seg 1/1 to: VIBCTXABA from: 1 to: 1372 1 AsnGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIlely 17

[illegible]

501 CCTCCGGACACATAGACGCTTGGAGGAGGAGACCGCGATTCATCATCACC 550
184 roGInGlycysGlyAsnSerSerArgThrIleThnGlyAspThrCysAsn 200
551 CCGCGGGTTTGTGGGATAGCTCCAAATATCATGATCAGTAAGTAACACTGGAT 600
201 GluGluThrGlnAsnLeuSerThrIleLeuArgGluIutyrgInserIy 217
601 GAAAAAACCCAAAGCTACGTCTTAAATTCCTTGACGATACCCAACTTAA 650
217 sValIysArgGlnIlePheSerAspTyrgInSerGlnValAspIleIy 234
651 AGTTAAAGACCAATATTTTCACGCTATCATTGATATTTGATATACATA 700
234 snArgIleArgAspGluLeu 240
701 ATGAAATTAAAGTATGATA 720

seq_name: gb_pat:AR118598

seq_documentation_block:

LOCUS	723 bp	DNA	linear	PAT 16-MAY-2001
DEFINITION	Sequence 7 from patent US 6149919.			

VERSION AR118598.1 GI:14100508

SOURCE	Unknown.
--------	----------

Unclassified

AUTHORS Domenighini, M., Rapp

LT, their preparation and their use for the preparation of the LT, and the use of the LT for the preparation of the LT.

FEATURES

Location/Qualifiers

```

/organism="unknown"
PAGE COLUMN 337 105 300 4

```

ORIGIN

```
alignment_scores:
```

Ratio:	4.814	Gaps:	0
--------	-------	-------	---

100

US-09-528-682-1 x AR118598

Align seg 1/1 to: AR118598 from: 1 to: 723

1 AsnG1AspArgTrpLeuTrpAlaGlnAspSerTrpProAspGluIleu 17
 11 |||:::|||||
 1 AAlaGAsnAlaGATGATTATATGGGCGACATTCTAGACCTCTGATGAATAAA 50
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 11 :::::|||||
 51 GCAGTCAGGAGGCTCTATGCCAAGAGGACAGAGTGAGTACTTGGCCGAG 100
 34 LyrHisGlnMetAsnIleAsnLeuTyrAspPheSAlaArgGlyThrGlnThr 50
 101 GTATCTCAATGATATCAACACTTTTATGTATCTGCAAGAGGAACTCAGAGC 150
 51 GlyPheValAlaGTrpAspGlyTyrValSerThrSerLeuSerLeuArg 67
 151 GGATTTGTAGAGCGACGATGATGAGATGTTCCACCTCAATATGATTGAG 200
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTrpTyr 84
 11 |||:::|||||
 201 AAGTGGCCACTTATAGTGGTCAACACTATATTGTCTGGCATCTTACTATTAT 250
 84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 1000

```

251 ATATATATGTTATAGCCACCTGCACCACCAACATGTTTAACTGTAATGATGTA 3000
      |||||
101 LeuGIyValIYrSerProHisProTYrGIuInGIuVaISerAlaLeuGI 117
      |||||
301 TTAGGGGCGATCAGCTCCATCCATCCAGATGAACACAGAAAGTTCCTGCTTAAAG 3500
      |||||
117 YGIYIleProYrSerGIuIleTYrGIYIprTYrArGIValAsnPhleGIY 134
      |||||
351 TGGCATTCATACATCCCAAAATATATGCGATGCTATCCAGATTCATTTGCGG 4000
      |||||
134 alIleAspGIuArGIeuhISArGsnArGIuTYrArGAspArGIYTYr 1500
      |||||
401 TGCTTGATGACAACTATACATGCTATATAGGGGCTTCACAGATACATATTAC 4500
      |||||
151 ArgAsnLeuAsnIIleAlaProAlaGIuAspGIYrTYrArgLeuAlaGIYpH 167
      |||||
451 AGTAACTTAGATATTGCTCCACGACGAGATGGTATGATGGACAGGTTT 5000
      |||||
167 eProPArGAspHISGIuAlaATrPArGIuGIuIuProTrPIleHISAlaIAP 184
      |||||
501 CCTCTGGAGACATAGAGCTTGAGAGGAAGCCGTGATTCATCATCATGCAC 5500
      |||||
184 roGIuInGIYcGIYAsnSerSerArGhrIleThrGIYAspThrCysAsn 200
      |||||
551 CGCCGGGTTGTGGGAATCTCCCAAGATCATCGATCACTAATACTTCGAT 6000
      |||||
201 GIuGIuThrGIuAsnIleuSerThrIleTYrIleuArGIuIuTYrGInSerIY 217
      |||||
601 GAAAAAACCCAAAGCTAGTGCTAAATTCCTTGACAAATACCATTCATA 6500
      |||||
217 sValIYsArGIuIlePhSerAspTYrGIuInSerGIuValAspIleTYrA 234
      |||||
651 AGTTAAAGACAAATATTTTTCAGCTCATCATCGATATGATATACACATA 7000
      |||||
701 ATGAGATTAAGGATGAATTA 720

```

seq_name: gb_pat:E00132

```
seq_documentation_block:
```

DEFINITION	DNA coding of cholera toxin.
100133	

VERSION E00132.1 GI:2168431

SOURCE	unidentified.
--------	---------------

unclassified

AUTHORS Arufuooru, N. and Do, B.M. ;
EDITOR CRIBBITE A AND B OF DNA RECOMBINATION CHOICEA TOYIN

TOURNAMENT AND MEDICINE
 Date sent: TD 10002222022-2 1 22-DEC-100022

SMITH KLINE RIT
OF scholars (vibrant)

PN JP 1983222033-A/1
FD 33-DEC-1983

PF 23-MAY-1983 JP 19
DP 24-MAY-1983 TS 82

PI ARUFUOORU NIJIERU, DO BUIRUDO MISHIERU
DC 261K38/106 261K38/00 C07H31/04 C13N15/00-3

CC strandedness: Double,
CC topology: Linear;

CC hypothetical: No;
CC anti-conso: No.

CC	*source:	strain=ATCC	39050;
FH	Location/Qualifications		

223

FT 1 77
E2+ 000+170

FT 774. .1148 /product='subunit A of cholera toxin' FT CDS
 FT mat_peptide 774. .1145 /product='subunit B of cholera toxin' FT
 FT Location/Qualifiers
 source 1. .1148
 /organism='unidentified'
 /db_xref='taxon:32644'
 BASE COUNT 382 a 184 c 235 g 347 t
 ORIGIN

alignment_scores:
 Quality: 1088.00 Length: 240
 Ratio: 4.814 Gaps: 0
 Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
 US-09-528-682-1 x E00132 ..

Align seg 1/1 to: E00132 from: 1 to: 1148

1 AsnGlyAspArgLeuYrArgAlaAspSerArgProProAspGluIleYr 17
 55 AATGATCATATAGTTATTCGGCGAGATTCTAGACTCTCTCATGCAATATAA 104
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 105 GCAGTCAGTGTCTATTCGCAAGAGACAGAGAGTACTTGCACCGAG 154
 34 lythrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 155 GTACTCAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGACG 204
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
 205 GGAATTTTGGCGACGATGATGATGTTCCACTCATAGATTGGAG 254
 67 gSerAlaHisLeuAlaGlyIleSerIleLeuSerGlyTyrSerThrTyrT 84
 255 AAGTCCCACTTAGTGGTCAACTATATGTCTGTCGTCATCTTATT 304
 84 yrlleTyrValIleAlaThrIleProAsnMetPheAsnValAsnAspVal 100
 305 ATATATATGTTATAGCCATGCGACCAACATGTTTAACTTAAGATGA 354
 101 LeuGlyValTyrSerProHisProTyrGluGlnIleValSerAlaLeuG 117
 355 TTAGGGGATACAGTCTCTCATCCAGATGACAGAGAGATTCTGCTTTAGG 404
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 405 TGGATTTCCATCTCCCAATATATGATGATGATGATGATGATGATGATG 454
 134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 455 TGGTGTATGACAACTTACATCGTATAGGGGCTACAGAGATAGATATTAC 504
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
 505 AGTACTTATATATGCTCCAGCGACAGATGTTATGATTTGGCAGGTTT 554
 167 eProProAspHisGlnAlaTyrArgGluGluProTyrIleHisAlaP 184
 555 CCGTCCGAGACATGAGCTTGGAGGAAGACCGTGGATTCATCATGACAC 604
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 605 CGCGGGGTTTGGGATGCTCAAGATCATGAGATGATTAATCTTGGCAT 654
 201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217

655 GAAAAAACCAAGCTAGGTGTAATAATTCCTTGACGATATCAATCTAA 704
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 705 AGTAAAGACAAATATTTTCAGGCTATCATCTGATTTGATACACATA 754
 234 snArgIleArgAspGluLeu 240
 755 ATAGCAATTAAGATGAATTA 774

seq_name: gb_ba:VCT0XAB21

seq_documentation_block:

LOCUS VCT0XAB21 1362 bp DNA linear BCT 23-SEP-1991
 DEFINITION V.cholerae genes for toxin proteins A and B, strain 2125.
 ACCESSION X58786
 VERSION X58786.1 GI:48420
 KEYWORDS cholera toxin.
 SOURCE Vibrio cholerae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 REFERENCE 1 (bases 1 to 1362)
 AUTHORS Dams,E.
 TITLE Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
 JOURNAL RUCA University Antwerp, Gruenendurgenlaan 171, B-2020 Antwerp,
 Belgium
 2 (bases 1 to 1362)
 AUTHORS Dams,E., De Wolf,M. and Dierick,W.
 TITLE Correction of the cholera toxin nucleotide sequence of the Vibrio
 cholerae eltor strain 2125

JOURNAL Unpublished
 FEATURES
 source Location/Qualifiers
 1. .1362
 /organism="Vibrio cholerae"
 /strain="eltor 2125"
 /db_xref="taxon:566"
 /clone="pRIT10824"
 72. .92
 /rpt_type="TANDEM"
 -35_signal 110. .121
 -10_signal 136. .142
 CDS 169. .945
 repeat_region
 /codon_start=1
 /transl_table=11
 /product="cholera toxin A protein (CTA)"
 /protein_id="CAA41592.1"
 /db_xref="GI:48421"
 /db_xref="SWISS-PROT:P01555"
 /translation="MKITFVFIFLSSFSYANDKLYRADSRPPEIKQSGGLPRG
 OSEYFDRTQNNINLYDHARGTQTEFVRHDDYVSTISLSAHLVGOTILSGHSTYY
 IYVATAPMFMVNDVLDGAYSPHDEQVSAVGGLPYSOITGMRYVHFEVLDEQLHRN
 RGYRDRYVNLIDIAPADYGLAGFPPEHRAAREPWIHAAPCGGNAPRRSSMSYTD
 EKTQSLGVFLELYQSKVKRQIFSGVQSDIDHNRIKDEL"
 169. .222
 223. .942
 /product="cholera toxin A protein (CTA)"
 942. .1004
 942. .1316
 /codon_start=1
 /transl_table=11
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41593.1"
 /db_xref="GI:48422"
 /db_xref="SWISS-PROT:P01556"
 /translation="MKLRFGEFVYLSSAVAHGNPONTDLCAEYHNTQYTLNDK
 IESYSESLAGKREMAIIPKNGAIFQVPEPSQHIDSOKKAIERKKDILRIYLTLEAK
 VERLCVWNKTPHAIATATSMAN"

mat_peptide 455 a 205 c 276 g 426 t
 BASE COUNT 455 a 205 c 276 g 426 t
 ORIGIN

alignment_scores:

Quality: 1088.00 Length: 240
 Ratio: 4.814 Gaps: 0
 Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x VIBCTXABB ..

Align seg 1/1 to: VIBCTXABB from: 1 to: 1369

```

1 AsnGLyAspArgLeuTyrArgAlaAspSerArgProPheAspGluLeu 17
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
164 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213
17 sarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
214 GCAGTCAGGTGGCTTATGCTCAAGAGACAGAGTACTGACCTGACCGAG 263
34 lYThrGlnMetAsnLeuAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
264 GATCTCAAAATGAATATCAACCTTATGATCATGCAAGAGCACTCAGACG 313
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
314 GATTTGTTAGGACAGATGATGATGATGATGATGATGATGATGATGATG 363
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
364 AAGTCCCACTTATGCTGCTCAAACTATGCTGCTGCTGCTGCTGCTGCT 413
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
414 ATATATATGTTATAGCCACTGCGACCCACACATGTTTAACGTTAATGAT 463
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
464 TTAGGGGACATACAGTCTCATCCAGATGAACAAGAGTTTCTGCTTAGG 513
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
514 TGGGATTCATCTACTCCCAATATATATGATGATGATGATGATGATGAT 563
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
564 TGGTGTATGATACATATACATCGTAATAGGGGCTACAGAGATGATATTA 613
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgGluAlaGly 167
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
614 AGTAATGATGATTTGCTCCAGACAGCATGCTTATGATGATGATGATGAT 663
167 eProProAsnHisGlnAlaIleTyrArgGluProTyrIleHisHisAla 184
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
664 CCTCCGGAGCATAGAGCTTGGAGGAGAGAGCCGCTGATTCATCATGCA 713
184 rGlnIleYsgIysAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
714 CGCCGGGTGTGGAAATGCTCCAAAGATCATGATGATGATGATGATGAT 763
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuAlaGluTyrGlnSer 217
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
764 GAAATAACCAAGCTAGTGTAAATTCCTTGACGAATACCAATCATTA 813
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
814 AGTTAAAGAACAAATATTTTCAAGCTATCAATCTGATGATGATGATGAT 863
234 snArgIleArgAspGluLeu 240
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
864 ATAGAAATTAAAGATGAATTA 883

```

seq_name: gb_ba: VCTOXA56

seq_documentation_block:

LOCUS VCTOXA56 1397 bp DNA linear BCT 30-JUN-1993
 DEFINITION V.cholerae genes for toxin proteins A and B, strain 569B.
 ACCESSION X58785 S53782
 VERSION X58785.1 GI:48888
 KEYWORDS cholera toxin.
 SOURCE
 ORGANISM Vibrio cholerae.
 Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 Dams, E.
 Dams, E. (bases 1 to 1397)
 Direct Submission
 Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
 RUCA University Antwerp, Gruenembergerlaan 171, B-2020 Antwerp,
 Belgium
 2 (bases 1 to 1397)
 Dams, E., De Wolf, M. and Dierick, W.
 Nucleotide sequence analysis of the CT operon of the Vibrio
 cholerae classical strain 569B
 JOURNAL Biochim. Biophys. Acta 1090 (1), 139-141 (1991)
 MEDLINE 91355224
 FEATURES
 source
 Location/Qualifiers
 1..1397
 /organism="Vibrio cholerae"
 /strain="classical strain 569B"
 /db_xref="taxon:666"
 /clone="PUM17"
 72..127
 /rpt_type="TANDEM
 145..156
 171..177
 204..980
 /codon_start=1
 /transl_table=11
 /product="cholera toxin A protein (CTA)"
 /protein_id="CAA41590.1"
 /db_xref="GI:48889"
 /db_xref="SWISS-PROT:P01555"
 /translation="MKITIEFFFLFSFSYANDDKLYRADSRPPDEIKQSGGLMPRG
 QSEYFDRTGNINILYDHARCTGTFVHDGIVSTISLSAHLYGOTILSGSTTY
 IYVATAVNMFNVLDGAVSPHDPDEVSALGIPYQIYGWRVHFGVLDQHLRN
 RGYDRYSNDIADPADGYGLAGFPPEHRAHREPPIYHAPPCGNAHRRSSMSWTC
 EKTSQIGYKFLFDEYQSKVKRQIFSGYQSDIDTNRIRKDEL"
 204..257
 258..977
 /product="cholera toxin A protein (CTA)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPTREMBL:057193"
 /translation="MIKKRGVFFVLLSSAYANGTPONTIDCAEYHNTQIHLNDK
 IFSTESLAGREMAIITFKNGATFQVEYPSQIHDSKKAIEKKDILRIAYLLEAK
 VEKICWNKTPHAIATSMAN"
 1040..1348
 /product="cholera toxin B protein (CTB)"
 /protein_id="CAA41591.1"
 /db_xref="GI:48890"
 /db_xref="SPT

570 AATGATGAATAGTATTATCGGGCAGATTCTAGACCCGCCGATGAATAA 619

17 SARFSGRGGLYLEUMETPROARGLYHISANGIUTYRPHASPARG 34
|||||.....:|||||

620 GCAGTCAGGTGCTTTATGCCAAGAGGACAGTAGTGACTTTGACCAG 669

34 LYTHGLIMETSNILEASNLEUYRASPNSIALARGLYLPHRGINTPR 50
|||||.....:|||||

670 GTACGCAAAATGAATACAACTTTATGATCATGCACAAGAGAACCTCAGAG 719

51 GLYPHEVALARYTRASPAPSLYTYSVALSERTHSERIESEURLEU 67
|||||.....:|||||

720 GGATTTGTTAGGCACGATGATGATATGTTTCCACCTCATATNGTTGAG 769

67 GSERALHISLEUALAGLYGINLERILEUSERGITYRSERTHYRT 84
|||||.....:|||||

770 AAGTCCCCCTAGTGGGTCCAACATATATGCTGGTCAATTCCTTAAT 819

84 YLLLETYRALLLEATHRALAPROASMETHESNVALSNAPVAL 100
|||||.....:|||||

820 ATATATATGTATATACCCACCTGCACCAACATGTCTTAATGATGTA 869

101 LEUGLYVALTYRSERPROHSIPROYTGINGLIVALSERALALEUG 117
|||||.....:|||||

870 TTAGGGGCGATACAGTCCTCATCCAGATGACAGAAGAGTTTCGCTTAG 919

117 YGLYLEPROTYRSERGINLIETYRGLYTPPYRARVALASNPHGLYV 134
|||||.....:|||||

920 TGGGATTCATCTCCCAAATATATGATGATGATCGATTCATTTTTGGG 969

134 ALLIASPOLUARGLUNHISARGASNARGLUITYRARGSPARTGYTRYR 150
|||||.....:|||||

970 TCCTTGATGAAACAATYACATYCCTAATAGGGCGCTACAGATAGATTTAC 1019

151 ARGASNLEUASNILLEALARPROLAGIUASPGLYTYRARGLEUALIGLYP 167
|||||.....:|||||

1020 AGTAACTTGATATTTGCTCCACAGACAGATGGTTATGAGATTGGCAGTTT 1069

167 EDROROPASPHISGLINALATRPARGLUINUPROTPIENSHISALAP 184
|||||.....:|||||

1070 CCGTCGCGAGCATACAGCTTGGAGGGAAGACGCGTGATTCATCATGAC 1119

184 RGLINGLYCYSGLYASNSESRERATGTHILETHCLYASPHRCYSAAN 200
|||||.....:|||||

1120 CGCCGGGTGTGGGAATGCTCCAAAGATCATCATCATATCTTCCGAT 1169

201 GIUGIUTHRGINSLEUSERTHRILETYRLIETRYLARAGLUITGLINSEIR 217
|||||.....:|||||

1170 GAAAAAACCCCAAGCTTAGGTTAAATATTCCTTGACGAATACCAATCTTA 1219

217 SVALYLSARGLINIEPSESERPYSRGINSERGIVUALASPILETYRA 234
|||||.....:|||||

1220 ACTTAAAAGACAATATTTTTCAGGCTATCATCAATCTGATATGATACACANA 1269

234 SNARGIILEARGPSGLIUDEN 240
|||||.....:|||||

1270 ATAGAATTAAAGATGAATTA 1289

seq_name: gb_pat:E03512

seq_documentation_block:

LOCUS E03512 2020 bp DNA linear PAT 29-SEP-1997

DEFINITION Cholera toxin gene.

VERSION E03512

KEYWORDS JP 1992099488-A/17.

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 2020)

AUTHORS Shiral,H.

TITLE PRIMER FOR AMPLIFICATION OF GENE

JOURNAL Patent: JP 1992099488-A 17 31-MAR-1992:

[illegible]


```

/transl_table=11
/product="zot"
/protein_id="AAL09684.1"
/db_xref="GI:15825400"
/translation="MSIFRHHGAPGSKYKTSGALWLRLLPAIKSGRIITVNGILNER
MAYLKMDVSDISIEFIEDIDHPDGLRMFAFMHARKDAPLIDECEGIMPRLTAN
LKALDPDPDLVEDRPESEFVADMDHMHQMDCLTPNPAKYHMTREAEIGYRHF
NRATVGGAKETLTTHDAASGQMSHALROYKRIIPSEIKMYASTTGKARDTMAG
TALKMDKRIIFLEGWYLMFSYFGYGLNDIPITGSDAITEESQSPQSKATAGNAV
GSKAPVASFGEFCIGRLCVDGFTVGDENRLDNDIDITREGMAIGHIYKDLTV
FFEFSSGVPTELEFASSYRYKVLPLDPENHVFDFPAALWVEARGLPIETENDK
KGNLSTP"
3128..3904
/gene="ctxA"
3128..3904
/gene="ctxB"
/codon_start=1
/transl_table=11
/product="ctxA"
/protein_id="AAL09681.1"
/db_xref="GI:15825397"
/translation="MKRIIFPFILFSPSYANDDKLYRADSRPDEIKOSGLMPG
QSEFDRGTQMNINLYDHARGTQTGPRHDDGYSTISLRSHLVQOTILSGHSTY
IYIATAPNMFNVNDVLGAYSPPDDEQVSALGIPYSQLYGWVHVGVLQDLHRN
RGVRDRYSNMDIAPADGVLGAGFPENHARBEPMIHAAPGCCNAPSMSMNC
EKYOSLCKEFLDEQVSKYRQIFSGYSDIDITHNRIKDEL"
3901..4275
/gene="ctxB"
3901..4275
/gene="ctxB"
/codon_start=1
/transl_table=11
/product="ctxB"
/protein_id="AAL09682.1"
/db_xref="GI:15825398"
/translation="MKRLKGFVEFTVYLSSAYAHGPHNITATCAEYHNQITLNDK
IFSTSLAGKREMAITFERNGATFOVEYVPSGHSIDSKAIEIRMDTLIAYLTEAK
VERKCVNNNTPHAIATISMAN"
BASE COUNT 1132 a 904 c 997 g 1242 t
ORIGIN

Alignment scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

Alignment block:
US-09-528-682-1 x AF414369 ..

Align seg 1/1 to: AF414369 from: 1 to: 4275

1 AsnglyAspargLeuTyrrgAlaAspSerArgProGloAspGluile 17
||| |||:::|||||
3182 AATGATGATAACTTATATCGGACAGATTCTACACCTCCGATGAATAA 3231

17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyIuTyrrPheasparG 34
1:::|||||:::|||||:::|||||
3232 GCAGTCAGGTGGCTTATGCCAAGAGGACAGAGTAGTACTTGGACCGAG 3281

34 IYThGlnMetAsnIleAsnLeuTyrrAspHisAlaArgGlyIhrGlnThr 50
|||||:::|||||:::|||||:::|||||
3332 GGATTGTAGGACAGATGATGATGTTTCACACTCAATTAAGTTTGAG 3381

51 GlyPheValArgTyrrAspAspGlyTyrrValSerThrSerLeuSerLeuAr 67
|||||:::|||||:::|||||:::|||||
3382 AAGTCCGCCCTAGTGGGTCAACATATATGCTGTCATCTTACTTATT 3431

84 YrIIETyrrAlIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100

```


|||||
3432 ATATATATGTTATGACACGACCCACCATGTTAACTTATGATGTA 3481
101 LeuG1ValTyrSerProH1SerProTyrGluGluValSerAlaLeuG1 117
3482 TTAGGGGACATAGCTCTCATCCAGATGAACAAGATTTCTCTTAGG 3531
117 yG1LePProTyrSerGlnIleTyrGlyTyrTyrValAsnPhcGly 134
3532 TGGGATTCCTACTCCCAATATATGATGATTCGATTCATTTTGGG 3581
134 a11LeaSPG1uaTgLeuH1aTgAsnTgG1uTyrTgAspTgTyrTyr 150
3582 TCGTTGATGAACAATTCATCATGTAATAGGGGCTACAGATAGATATATAC 3631
151 ArgAsnLeuSnIleAlaProAlaGluAspGlyTyrTgTgLeuAlaGlyPh 167
3632 AGTAACTTAATATGCTCCAGACGACGATGTTATGATTTGCGAGTTT 3681
167 eProProAsPH1sGlnAlaTTrpArgGluGluProTPrIleH1sH1a1aP 184
3682 CCCTCCGACATAGACCTTGGAGGGAAGACCGTGCATTCATCATGAC 3731
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspPhrCysAsn 200
3732 CGCCGGGTGTGGGAATGCTCAAGATCATGATGATTAATCTGCGAT 3781
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
3782 GAAAAAACCCAAAGCTAGGTGTAATAATCTTGCAGAAATCAACATCTAA 3831
217 sValTyrArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
3832 AGTTAAAGACAATATTTTCAGGCTATCATCTGATATGATACACATA 3881
234 snArgIleArgAspGluLeu 240
3882 ATGAATTAAGATGAATTA 3901
seq_name: gb_AE004224
seq_documentation_block:
LOCUS AE004224 10646 bp DNA linear BCT 31-JUL-2000
DEFINITION Vibrio cholerae chromosome I, section 132 of the complete
chromosome.
ACCESSION AE004224 AE003852
VERSION AE004224.1 GI:9655952
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
AUTHORS
1 (bases 1 to 10646)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermlaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uterback, T., Fleishmann, R.D.,
Nierman, W.C. and White, O.
TITLE
DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
JOURNAL
MEDLINE
REFERENCE
AUTHORS
2 (bases 1 to 10646)
20406833
Nature 406 (6795), 477-483 (2000)

FEATURES
source

Location/Qualifiers
1..10646
/organism="Vibrio cholerae"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:666"
/chromosome="I"
/note="biotype: B1 for"
complement(215..439)
/gene="VC1452"
complement(215..439)
/note="similar to GP:2564354; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstC protein"
/protein_id="AAF94609.1"
/db_xref="GI:9655953"
/translation="MSLKPYLMDVDSLIEDLNMMALYLSGAYTDEIAHQVNLICD
KIIDLQIVNFIKRLSPSINPOLKALTEPSL"
complement(533..907)
/gene="VC1453"
complement(533..907)
/note="similar to GP:2564353; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstB1 protein"
/protein_id="AAF94610.1"
/db_xref="GI:9655954"
/translation="MMKLMVINMKSRIYVFGASHSEGVSKTGAFLIPLVFGKPIRQ
WKNKGGLTGLQHOEKYKPVSSDAMVRKLEQTAEPVLVTFEDNEPDEDSRNIVIDY
QVCSLRPNVPGKRLDKPKKI"
complement(885..1964)
/gene="VC1454"
complement(885..1964)
/note="similar to GP:2564352; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstA1 protein"
/protein_id="AAF94611.1"
/db_xref="GI:9655955"
/translation="MKKQIFLDELQIDTNSPPVFDYLAWSVPVAFRRHAKSDLS
SLINAPLPKPYRKARPEQEKIIEIAKKQNNYAMMRLEVRCLHYLGIRMSPPRK
GLYGENSCHLMKSYSNKRVGFVALDGGKRNCTFQIBGVGRTYLEHTSLFRLMWLD
LLGCSRLSRIDLAVDDFGLRGREYAKAVSDDAFRTARAGRAPNGERLVSFNGKI
INESFEVGSRESRIYWRILYKRAQGLDMHFRNEVELKMPIDIVLLINIEGYPAGLCA
YSASTINSLPVKVVTKKROVALDIHSRKKMARQVGTLEFDSKHFGDLERVGALIS
KEIHDSINLPSDYMKIDELMD"
2090..2428
/gene="VC1455"
2090..2428
/note="similar to GP:2564351; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="transcriptional repressor RstR"
/protein_id="AAF94612.1"
/db_xref="GI:9655956"
/translation="MKIKERLANRKAIINKTQAQMADEIGLSLYKKYEGEGPLPM
ENVIKRIADALISIDELCGRRATDENELMKRLKIQQLDEEBQKATISMVLESLINH
STKSLNIGA"
complement(3030..3404)
/gene="VC1456"
complement(3030..3404)
/note="similar to GB:K01170 SP:P01556 GB:ID30052 GB:ID30053
GB:X58785; identified by sequence similarity: putative"


```

167 eProPAspHisGlnAlaIatPArgGluGluProTrpIleHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3623 CCCGCCGAGCATAGACTTGTGAGGAAAGACCCTGGATTCATCATGCAC 3574
184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3573 CGCCGGGTGTGGGAATGCTCCAAAGATCATGAGTAAATATCTGGCAT 3524
201 GluGluTrpHisAsnLeuSerThrIleThrLeuArgGluTrpGlnInserty 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3523 GAAAGAACCCAAAGCTAGGTAAATCTCTTACGATACCAATCAATCAA 3474
217 sValIysArgGlnIlePheSerAspTrpGlnSerGluValAspIleTry 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3473 AGTAAAGACAAATATTTTCAGGCTATCATCATCATTCATACACATA 3424
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3423 ATAGATTAAGATGAATTA 3404

seq_name: gb_ba:VIBCTXA
seq_documentation_block:
LOCUS VIBCTXA 777 bp DNA linear BCT 26-APR-1993
DEFINITION Vibrio cholerae enterotoxin A1 peptide (ctxA) gene, 5' end.
ACCESSION K02679
VERSION K02679.1 GI:155159
KEYWORDS ADP-ribosyltransferase; enterotoxin.
SOURCE V. cholerae (El Tor strain 62746) DNA.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
AUTHORS Lockman,H.A., Galen,J.E. and Kaper,J.B.
TITLE Vibrio cholerae enterotoxin genes: Nucleotide sequence analysis of
JOURNAL DNA encoding ADP-ribosyltransferase
MEDLINE 85006737
COMMENT This sequence is reported [1] to be identical to that for strains
El Tor 2125 and 569B. the 'titigat' sequence repeated four times
beginning at position 38 and once more at 101 is correlated with
the amount of enterotoxin produced in various strains.
FEATURES
source
location/Qualifiers
1..777
/organism="Vibrio cholerae"
/db_xref="taxon:666"
142..195
/feature="enterotoxin signal peptide"
142..>777
/feature="enterotoxin prepeptide"
/feature="start=1"
/feature="table=11"
/feature="id="AA27514.1"
/db_xref="GI:155160"
/feature="translaction="MWKILFVFIFLISFSYANDKLYADSRPDEIKOSGGLMPRG
GSEYFDRTQMININILDHARGTGTGTHDDGVSTISLSRSHLVGOTILGSHSTYV
IVYIATAPNMFNVDYLGAYSPHDEQESAGLIPYSOIGWYRVHFGVLEQLHRN
RGYRDYYSNLDIAPADAGYGLAFPEPRHAWREEMHHPGCCNADRSS"
196..777
/product="enterotoxin subunit A1"
BASE COUNT 237 a 119 c 164 g 257 t
ORIGIN 189 bp upstream of NdeI site.

alignment_scores:
Quality: 934.00 Length: 193
Ratio: 5.049 Gaps: 0
Percent Similarity: 95.585 Percent Identity: 86.528

alignment_block:
US-09-528-682-1 x VIBCTXA ..
Align seg 1/1 to: VIBCTXA from: 1 to: 777

```

```

1 AsnGlyAspArgLeuTrpArgAlaAspSerArgProProAspGluIlely 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 AATCATATATTAAGTATATCGGCGAGATTTAGACCTCTCATGAATAAA 245
17 sarGSerGlyGlyLeuMetProArgGlyHisAsnGluTrpPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 GCACTCAGAGGTCTTATGCGCAAGAGACAGAGAGTACTTTCAGCCAG 295
34 IyThrGlnMetAsnIleAsnLeuTrpAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 GTACTCAAAAGAAATATACCTTATGATCATGCAAGAGAACTCAGACG 345
51 GlYpHeValArgTrpAspAspGlyTrpValSerThrSerLeuSerLeuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 GGATTTTTCAGCGACGATGATGATGTTCCACCTCATTTAGTTGAG 395
446 ATATATATGTTATAGCCAGTCCACCAACATGTTTAACTTAATGATGA 495
84 YrIleTrpValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 LeuGlyValTrpSerProHisProTrpGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 TTAGGGGCATACACTCTCATCCAGATGACAGAAAGAACTTCGTTTAG 545
117 yGlyIleProTrpSerGlnIleTrpGlyTrpTrpArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
546 TGGGATTCACATCTCCCAATATATGATGATGATGATGATGATGATG 595
134 alIleAspGluArgLeuHisArgAsnArgGluTrpArgAspArgTrpTr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
596 TCGTGTATGACATATCATGATGATGATGATGATGATGATGATGATG 645
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTrpArgLeuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
646 AGTAACTTAGATATATGCTCCAGCAGCAGATGTTATGATGAGAGTTT 695
167 eProPAspHisGlnAlaIatPArgGluGluProTrpIleHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 CCCGCCGAGCATAGACTTGTGAGGAAAGACCCTGGATTCATCATGCAC 745
184 roGInGlyCysGlyAsnSerSerArgThr 193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
746 CGCCGGGTGTGGGAATGCTCCAAAGATCA 774

seq_name: gb_ba:ECOEToxHL
seq_documentation_block:
LOCUS ECOETOXHL 1262 bp DNA linear BCT 18-NOV-1994
DEFINITION Escherichia coli (strain 41) heat-labile enterotoxin type IIB
(LT-IIB) A and B chain genes, complete cds.
ACCESSION M28523
VERSION M28523.1 GI:576584
KEYWORDS LT-IIB gene; enterotoxin type IIB.
SOURCE Escherichia coli (strain 41) DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS Pickett,C.L., Twiddy,E.M., Coker,C. and Holmes,R.K.
TITLE Cloning, nucleotide sequence, and hybridization studies of the type
JOURNAL IIB heat-labile enterotoxin gene of Escherichia coli
MEDLINE 89359131
COMMENT On Nov 28, 1994 this sequence version replaced gi:341953.
FEATURES
source
location/Qualifiers
1..1262
/organism="Escherichia coli"
/strain="41"

```



```

-35_signal      /db_xref="taxon:562"
                  /clone="pcr4185"
gene            1..112
                  /gene="LT-IIIb"
                  1..1262
                  /gene="LT-IIIb"
-10_signal      26..31
                  /gene="LT-IIIb"
                  113..904
                  /gene="LT-IIIb"
CDS            /note="A chain of heat-labile enterotoxin type IIB"
                  /codon_start=1
                  /transl_table=11
                  /product="enterotoxin"
                  /protein_id="AA53285.1"
                  /db_xref="GI:576585"
                  /translation="MAKVISFSLIFSLFPLVANDYFRADSRTPDEVRRSGGILPRG
ODAVEEGPFINILNDYHARGTATGNTNRNDGVSTTTLROAHLGONNLGYNERY
IYVAAAPFDYNGVIGRYSPYSENEYAAGCIPISQITGWRYSFGAIEGSMHNN
RDRRLDFRLSAAPNEDGYRIAGFPDGFVAMEVPRFRAPNSCLPNNKASSDTTCA
SLNKLSSQHDLAFFKTKIRKFTLMTLLSINNGFFSNGKDEL"
sig_peptide     113..172
mat_peptide      /gene="LT-IIIb"
                  173..901
                  /gene="LT-IIIb"
                  /note="A chain of heat-labile enterotoxin type IIB"
                  /product="enterotoxin"
sig_peptide      894..962
                  /gene="LT-IIIb"
                  894..1262
                  /gene="LT-IIIb"
                  /note="B chain of heat-labile enterotoxin type IIB"
                  /codon_start=1
                  /transl_table=11
                  /product="enterotoxin"
                  /protein_id="AA53286.1"
                  /db_xref="GI:576586"
                  /translation="MSFKIITKAFVIMAAVLSQAHAGASQFKNCRNTTASIVGCV
ELRKYSIDINNTDGMVYVSSGTGWRISRAKDPDMVMTAEMRKIMAAVLSGMRYVNE
MCASPSPNVIVALEAE"
                  963..1259
                  /gene="LT-IIIb"
                  /note="A chain of heat-labile enterotoxin type IIB"
                  /product="enterotoxin"
BASE COUNT      369 a 195 c 285 g 413 t
ORIGIN
Alignment_scores:
      Quality: 662.00      Length: 216
      Ratio: 3.875      Gaps: 0
Percent Similarity: 81.481      Percent identity: 57.407

alignment_block:
US-09-528-682-1 x ECOETOXHL ..
Align seg 1/1 to: ECOETOXHL from: 1 to: 1262

6 TYRARGALASPSEARAPROPROASGUILLELYASRSGLYLYE 22
|||||
182 TTCAGGCGCTGATTCGAGAACGCTGTGTAAGTCACAGCGCTCAGGGGGGTT 231
|||||
22 UMETPROARGLYNHLSASNGLUITYRPHESAPARFGLYTRGIMETASNI 39
|||||
232 GATTCACAGAGGTCAGGATGATGAGGCTTATGACGTGCGACGCCATTAAACA 281
|||||
39 IASNSLEUTYRASPHISALAAARGLYTHGILTHRGILPHEVALARGTYR 55
|||||
282 TTAACTATATNTGACCATGCACGAGACGCAACAAGGGAACACTAGATAT 331
|||||
56 AAPAARGLYTYRVALSERTHRSELEUSERTLEUARGSERALNHISLEUHL 72
|||||
332 AATGATGATATGTATCAACAACACGACTTGTGAGACAGGCTTACTTATTT 381
|||||

```

```

72 aclyInserIleLeuSerGlyTyrSerThrYrYrIleValIleA 89
   |||||.....|||||.....:|||||.....:
382 AGGGCAAAATATGCCTTGCGGTATAATAGACRATATTATGTTGTCG 431
      89 laThraAlaProAsnMetPheAsnValAsnAspValLleuGIyValTyrSer 105
         |||||.....:|||||.....:|||||.....:
432 CTGCAGCACCAAAATTGTTGATGTAAATGGCGCTTTTAGCAGAGTACT 481
      106 ProHISProTYrGLUGlnGluValSerAlaLeuGlyIleProTYrSe 122
          |||||.....:|||||.....:|||||.....:
482 CCGTATCCCAAGTGAATAATGATGCTGCATTAGAGGTATGCCCTGTC 531
      122 rGlInIeTYrGLTYrPYrArGVaIAsnPhcGIyValIeAspGIuArgL 139
          |||||.....:|||||.....:|||||.....:
532 ACAAATAATGGCTGCTATAGAGTATCTTCGCTGCTATAGAGGGGGGAA 581
      139 euHIstArgSnaTrgLUrTYrARgSPArGYrTYrARGasnuLeuAnIle 155
          ::::|.....:|||||.....:|||||.....:
582 TGCATCGAAACAAGCGATACAGAGAAGATTATTAGAGCGTTATGCT 631
      156 AlAProAlaGIuAspGLyTYrArGVaIeAlaGlyPheProProAspHisel 172
          |||||.....:|||||.....:|||||.....:
632 GCTCCTAAAGAGGATGCGCTATCGAAATGCTGGAATTCGCCGAGGATTTCC 681
      632 TGCGTGGGAGAGATTCCTCGAGAGAAATTTGCGCTTAACCTTGCTAC 731
          |||||.....:|||||.....:|||||.....:
172 nAlarTPArGVaIGLuPrOTripriLleHISAlaProGlnGIyCySGlyA 189
          |||||.....:|||||.....:|||||.....:
682 TGCGTGGGAGAGATTCCTCGAGAGAAATTTGCGCTTAACCTTGCTAC 731
      189 snSerSerArGThrIleThrgIyAsPTpHCysasnGIuThrgIasn 205
          ::::|.....:|||||.....:|||||.....:
732 CAATATATAAGCTGCCAGTACTACCTGCGCTCTTAAACAATAA 781
      206 LeuSerThrIleTYrLeuArgGIuTYrGlnSerLySvalLYaArgGln 221
          |||||.....:|||||.....:|||||.....:
782 CATACACAGCATGATTACCTGATTTTAAGAAATATATPAAGAGAAA 829
seq_name= gb_ba:ECOLTRI1A

seq_documentation_block:
LOCUS           ECOLTRI1A                1200 bp        DNA             linear       BCT_26-APR-1993
DEFINITION     E.coli heat-labile enterotoxin type IIA (LT-IIa) A and B genes,
complete cds.
ACCESSION     M17894
VERSION       M17894.1 GI:146671
KEYWORDS      enterotoxin; heat-labile enterotoxin.
SOURCE        E.coli (strain SA53) DNA.
ORGANISM      Escherichia coli
               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
               Escherichia.
REFERENCE     1 (bases 1 to 1200)
AUTHORS      Pickett,C.L., Weinstein,D.L. and Holmes,R.K.
TITLE         Genetics of type IIA heat-labile enterotoxin of Escherichia coli:
operon fusions, nucleotide sequence, and hybridization studies
JOURNAL      J. Bacteriol. 169, 5180-5187 (1987)
MEDLINE      880332841
FEATURES
    source
        location/Qualifiers
            1..1200
            /organism="Escherichia coli"
            /db_xref="taxon:562"
            60..839
            /note="heat labile enterotoxin type IIA A"
            /codon_start=1
            /transl_table=11
            /protein_id="AAA24093.1"
            /db_xref="GI:146672"
            /translation="MKHVLLFEVIFSVASANDFRADSRTPDEIRRAGGLPRGOO
EAYGRTPINILEHARGVTGNTNRINDGVSTTVLRQAHILGONILGSYNBYIY
VVAAPNLEFNGVLGYRGYSRSENEFAALAGLISDIIIGIVRSFGAIEGQMRFND
VRGDLFGTLVAPREDGYLAGPSNSNPAMEKMWSTFAPOCCVPNNKEFGVCISAN
TNVLSKDLMNFKKLLRRRLALTFRFSNEDDEIGVHGREDL."
            829..1200
            /note="heat labile enterotoxin type IIA B"

```


BASE COUNT 286 a 142 c 153 g 193 t
ORIGIN

alignment_scores:

Quality: 357.00 Length: 68
Ratio: 5.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 95.588

alignment_block:
US-09-528-682-1 x ECOELT ..

Align seg 1/1 to: ECOELT from: 1 to: 774

```

173 AAlATPARGlUGlUPrOTPrIleHISAlAPrOGInGIyCSGIyAS 189
|||||
2 GCTTGAGAGAGAAACCTGATTCATCATGCACACAGCTGTGGAGA 51
|||||
189 nSerSerArgThrIleThrgIyAspThrCysAsnGlUGlUThrgInAsnL 206
|||||
52 TTTCATCAAGAACATTTACAGCTGATCTTGTAAATGAGAGACCCAGAAATC 101
|||||
206 euserThrlIeTyrlleuArglUThrgInserLySvallySArgGlnIle 222
|||||
102 TGACACAAATATATCTCAGAAATATCATCAAAAGTTAAAGACGACATA 151
|||||
223 PheSerAspTyrgInserGluValAspIleTyrsnArgIleArgAspGI 239
|||||
152 TTTTCAGACTATCAGTCAGAGTTGACATATATTAACAGAAATTCGGAATGA 201
|||||
239 uLeu 240
|||||
202 ATTA 205

```

seq_name: gb_ba:ECLT87

seq_documentation_block:

LOCUS ECLT87 595 bp DNA linear BCT 28-OCT-1996
DEFINITION E.coli LT87 gene for heat-labile enterotoxin.
ACCESSION X83966
VERSION X83966.1 GI:1648865
KEYWORDS heat-labile enterotoxin; LT 87 gene.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE 1 (bases 1 to 595)

AUTHORS Germani,Y. and Desperrier,J.M.

TITLE Nucleotide sequence variations in genes encoding heat labile enterotoxins of Escherichia coli isolated in South Pacific

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 595)

AUTHORS Germani,Y.

TITLE Direct Submission

JOURNAL Submitted (17-JAN-1995) Y. Germani, Institut Pasteur, 28 rue du Dr

COMMENT Roux, F-75724 Paris Cedex 15, FRANCE

FEATURES Related sequence: M15361-3 (Yakamoto).

location/Qualifiers

1..595

/organism="Escherichia coli"

/strain="ETEC LT 87"

/db_xref="taxon:562"

205..579

/gene="LT 87"

205..579

/gene="LT 87"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin"

/protein_id="CA53880.1"

/db_xref="GI:1648865"

/db_xref="SWISS-PROT:P13811"

/translation="MNKVKFYLFETALLSLCAHGAAPQSTELCESEYHNTQIYTTINDK

BASE COUNT 226 a 105 c 113 g 151 t
ORIGIN
ILSYTESMAGKREMYITPEKSGATFOVEVGSQIHDSOKKAIERMKDPLRLTYLTETK
IDKICWNNKPTPNSIAISMEN"

alignment_scores:

Quality: 340.00 Length: 68
Ratio: 5.075 Gaps: 0
Percent Similarity: 98.529 Percent Identity: 97.059

alignment_block:
US-09-528-682-1 x ECLT87 ..

Align seg 1/1 to: ECLT87 from: 1 to: 595

```

173 AAlATPARGlUGlUPrOTPrIleHISAlAPrOGInGIyCSGIyAS 189
|||||
3 GCTTGAGAGAGAAACCTGATTCATCATGCACACAGCTGTGGAAA 52
|||||
189 nSerSerArgThrIleThrgIyAspThrCysAsnGlUGlUThrgInAsnL 206
|||||
53 TTTCATCAAGAACATTTACAGCTGATCTTGTAAATGAGAGACCCAGAAATC 102
|||||
206 euserThrlIeTyrlleuArglUThrgInserLySvallySArgGlnIle 222
|||||
103 TGAG.ACAATATATCTCAGAAATATCATCAAAAGTTAAAGACGACATA 151
|||||
223 PheSerAspTyrgInserGluValAspIleTyrsnArgIleArgAspGI 239
|||||
152 TTTTCAGACTATCAGTCAGAGTTGACATATATTAACAGAAATTCGGAATGA 201
|||||
239 uLeu 240
|||||
202 ATTA 205

```

seq_name: gb_ba:ECOELTA2

seq_documentation_block:

LOCUS ECOELTA2 165 bp DNA linear BCT 26-APR-1993
DEFINITION E.coli heat labile enterotoxin subunit a gene; a1/a2 junction.
ACCESSION K00433
VERSION K00433.1 GI:145829
KEYWORDS enterotoxin.
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE 1 (bases 1 to 165)

AUTHORS Spicer,E.K., Kavanaugh,W.M., Dallas,W.S., Falkow,S.,

Konigsberg,W.H. and Schaefer,D.E.

TITLE Sequence homologies between a subunits of Escherichia coli and

JOURNAL Vibrio cholerae enterotoxins

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 78, 50-54 (1981)

COMMENT 81223767

FEATURES see other loci beginning <ecoeelt>.

location/Qualifiers

1..165

/organism="Escherichia coli"

/db_xref="taxon:562"

59 a 30 c 35 g 41 t

about 300 bp after seg 1.

alignment_scores:

Quality: 247.00 Length: 53
Ratio: 4.843 Gaps: 1
Percent Similarity: 96.226 Percent Identity: 90.566

alignment_block:
US-09-528-682-1 x ECOELTA2 ..

Align seg 1/1 to: ECOELTA2 from: 1 to: 165


```

179  TrrlleHishisAlaProGlnGlyCysGlnSerSerArgThrIleTh 195
|||||
7  TGAATTCATCATGCACCAAGGTTGTGAGATTCATCAACAACATTCAC 56
195  rGlyAsPthrcysAsngluGluThrGlnAsnLeuSerThrIleTyrllea 212
|||||
57  AGGTCATCTACTGTATATGAGAGACCCAGAACCTGAGCAAAATATATCTCA 106
212  rGgluTyrglnSerIysValIys  ArgGlnIlePheSerAspTyrglnSe 228
|||||
107  GGGATATCAATCAAAAGTATTAGAGGCACATATTTTGACTATCAGCTC 156
228  rGluVal 230
|||||
157  AGAGCTT 163

seq_name: gb_ba:VCCTXAB2

seq_documentation_block:
LOCUS      VCCTXAB2                861 bp    DNA       linear    BCT 29-JUL-1994
DEFINITION V.cholerae (0139) ctxA and ctxB genes, 861bp.
ACCESSION  X76391
VERSION    X76391.1 GI:433859
KEYWORDS   cholera toxin; CTx2 protein; CTB protein; ctxA gene; ctxB gene.
SOURCE     Vibrio cholerae.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE  1 (bases 1 to 861)
AUTHORS    Lebens,M.
TITLE      Direct Submission
JOURNAL    Submitted (24-NOV-1993) M. Lebens, Dept. of Medical Microbiol. &
MEDLINE    Immunol., Goetegorgs univ., Guldhedsatan, Ioa, 413 46 Goeteborg,
FEATURES    SWEDEN
            location/Qualifiers
            1..861
            /organism="Vibrio cholerae"
            /strain="4260B"
            /isolate="serotype 0139"
            /db_xref="taxon:666"
            37..177
            /gene="ctxA2"
            /gene="ctxA2"
            /codon_start=1
            /transl_table=11
            /product="cholera toxin A2"
            /protein_id="CA53975.1"
            /db_xref="GI:433860"
            /db_xref="SPTREMBL:O57372"
            /translation="MSNTCDERTQSLGVKFLDEYQSKVKRYFSQYSDIDTHNRKID
            EL"
            174..548
            /gene="ctxB"
            /product="cholera toxin B subunit"
            174..548
            /gene="ctxB"
            174..548
            /gene="ctxB"
            /codon_start=1
            /transl_table=11
            /product="cholera toxin B subunit"
            /protein_id="CA53976.1"
            /db_xref="GI:433861"
            /db_xref="SWISS-PROT:P01556"
            /translation="MIKLKRGVFFTYLLSSAYAHGTPQNTIDLCATYHNTQIYTLNDK
            ITSYTESLAGKREMAITTFKNGAIFQVEVPGSOHDSOKKAIERKMDTLRIAYLFEAK

```

```

BASE COUNT      306 a      139 c      150 g      266 t
ORIGIN
VEKICVNNKTPHAIATISMAN"
alignment_scores:
  quality: 182.00      Length: 55
  Ratio: 3.872      Gaps: 0
  Percent Similarity: 85.455      Percent Identity: 60.000
alignment_block:
US-09-528-682-1 x VCCTXAB2 ..
Align seq 1/1 to: VCCTXAB2 from: 1 to: 861
186  G|C|Y|G|L|V|A|N|S|E|S|E|R|A|P|T|H|I|E|T|H|G|L|Y|A|S|P|T|H|C|Y|S|A|N|G|L|U|
|||||
10  G|T|G|T|G|G|A|N|T|G|C|T|C|C|A|G|A|T|C|T|G|A|G|A|T|G|A|T|A|T|C|T|G|C|A|T|G|A|A|A|
60  A|A|C|C|A|A|G|C|T|A|G|C|T|A|A|A|T|C|C|T|G|A|G|A|T|A|C|C|A|T|C|T|A|A|G|T|T|A|
202  u|T|H|G|L|A|N|L|E|U|S|E|R|H|I|E|T|H|E|U|A|G|L|U|T|Y|G|L|N|S|E|R|Y|S|V|A|L|
|||||
60  A|A|C|C|A|A|G|C|T|A|G|C|T|A|A|A|T|C|C|T|G|A|G|A|T|A|C|C|A|T|C|T|A|A|G|T|T|A|
219  y|S|A|R|G|L|I|L|P|H|E|S|E|R|A|S|P|T|Y|G|L|N|S|E|R|G|L|U|A|L|A|S|P|I|E|T|Y|R|A|N|A|R|G|
|||||
110  A|A|A|G|A|C|A|T|A|T|T|T|T|C|A|G|G|C|T|A|C|A|T|G|A|T|A|T|G|T|A|C|A|C|A|T|A|A|T|A|G|A|
236  I|L|E|A|R|G|A|S|P|G|L|U|E|U| 240
|||||
160  A|T|T|A|G|A|T|G|A|T|T|A| 174

seq_name: gb_ba:VCCTXAB

seq_documentation_block:
LOCUS      VCCTXAB                938 bp    DNA       linear    BCT 10-FEB-1999
DEFINITION V.cholerae (0139) ctxA and ctxB genes, 938bp.
ACCESSION  X76390
VERSION    X76390.1 GI:433856
KEYWORDS   cholera toxin; CTx2 protein; CTB protein; ctxA gene; ctxB gene.
SOURCE     Vibrio cholerae.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE  1 (bases 1 to 938)
AUTHORS    Lebens,M.
TITLE      Direct Submission
JOURNAL    Submitted (24-NOV-1993) M. Lebens, Dept. of Medical Microbiol. &
MEDLINE    Immunol., Goetegorgs univ., Guldhedsatan, Ioa, 413 46 Goeteborg,
FEATURES    SWEDEN
            location/Qualifiers
            1..938
            /organism="Vibrio cholerae"
            /strain="4260B"
            /isolate="serotype 0139"
            /db_xref="taxon:666"
            37..177
            /gene="ctxA"
            /gene="ctxA"
            /codon_start=1
            /transl_table=11
            /product="CTx2 cholera toxin A2 subunit"
            /protein_id="CA53974.1"
            /db_xref="GI:433858"
            /db_xref="SPTREMBL:O57372"
            /translation="MSNTCDERTQSLGVKFLDEYQSKVKRYFSQYSDIDTHNRKID
            EL"
            174..548
            /gene="ctxB"
            /product="cholera toxin B subunit"
            174..548
            /gene="ctxB"
            /codon_start=1
            /transl_table=11
            /product="cholera toxin B subunit"
            /protein_id="CA53976.1"
            /db_xref="GI:433861"
            /db_xref="SWISS-PROT:P01556"
            /translation="MIKLKRGVFFTYLLSSAYAHGTPQNTIDLCATYHNTQIYTLNDK
            ITSYTESLAGKREMAITTFKNGAIFQVEVPGSOHDSOKKAIERKMDTLRIAYLFEAK

```


gene /gene="ctxb"
174..548
/gene="ctxb"
CDS 174..548
/gene="ctxb"
/codon_start=1
/transl_table=11
/product="CTB cholera toxin B subunit"
/protein_id="CA53973.1"
/db_xref="GI:433857"
/db_xref="SWISS-PROT:P01556"
/translation="MIKLFQVFFVLLSAYAGHPONTDCAEYHNTOYLTNDK
IFSYESLAGKREMAITFEKNGAIFQVEPQSHIDSKKAIERKKDLRIAYLLEAK
VEKLCVMNNKTPHAIATISMAN"
mat_peptide 237..545
/gene="ctxb"
/product="CTB cholera toxin B subunit"

BASE COUNT 330 a 151 c 165 g 292 t
ORIGIN

alignment_scores:
Quality: 182.00 Length: 55
Ratio: 3.872 Gaps: 0
Percent Similarity: 85.455 Percent Identity: 60.000

alignment_block:

US-09-528-682-1 x VCCTXAB ..

Align seg 1/1 to: VCCTXAB from: 1 to: 938

186 GLYSGIYASNSERARGTHRIETHRGLYASPTHRCYSANGLUIGL 202
|||||
10 GGTGTGGATGCTCCAGATCATCGATGATACCTGCGATGAAA 59
202 urhglnasnserserargthriethrglyaspthrcysanlgluigl 219
:|||||
60 ACCCAAGCTAGCTGCTAAATTCCTTGACGAATACCAATCTAAAGTTA 109
219 ysargglnlepheserargthriethrglyaspthrcysanlgluigl 235
|||||
110 AAGACAAATTTTTCAGCTATCATGTATGATACATATATAGA 159
236 llaargspgluieu 240
|||||
160 ATTAAGATGATTA 174

seq_name: gb_pat:A02701

seq_documentation_block:

LOCUS A02701 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production.
ACCESSION A02701
VERSION A02701.1 GI:344645

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 795).

AUTHORS

TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F

JOURNAL Patent: WO 8302456-A 1 21-JUL-1983;

FEATURES Location/Qualifiers

SOURCE

1..795 /organism="synthetic construct"

misc_feature

1..795 /db_xref="taxon:32630"

BASE COUNT 207 a 180 c 193 g 174 t

ORIGIN

alignment_scores:

Quality: 166.00 Length: 43
Ratio: 4.486 Gaps: 0
Percent Similarity: 86.047 Percent Identity: 79.070

alignment_block:

US-09-528-682-1 x A02701 ..

Align seg 1/1 to: A02701 from: 1 to: 795

188 GLYASNSERARGTHRIETHRGLYASPTHRCYSANGLUIGLURHGI 204
|||||
664 GGAATTCATCAAGACATCAAGAGCTGACTGTATGAGAGACCA 713
204 naslnserthrliethrglyaspthrcysanlgluiglyarg 221
|||||
714 GAATCTGACCAATATATCTCAGGGAATATCAATCAAAAGTTAAGGC 763
221 lnlepheserargthriethrglyaspthrcysanlgluigluval 230
|||||
764 AGATATTTTCAGACTATCAGCTCAAGSTN 792

seq_name: gb_pat:A02702

seq_documentation_block:

LOCUS A02702 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production (reverse
complement).

ACCESSION A02702

VERSION A02702.1 GI:344646

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 795)

AUTHORS

TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F

JOURNAL Patent: WO 8302456-A 2 21-JUL-1983;

FEATURES Location/Qualifiers

SOURCE

1..795 /organism="synthetic construct"

misc_feature 1..795 /db_xref="taxon:32630"

BASE COUNT 174 a 193 c 180 g 207 t

ORIGIN

alignment_scores:

Quality: 166.00 Length: 43
Ratio: 4.486 Gaps: 0
Percent Similarity: 86.047 Percent Identity: 79.070

alignment_block:

US-09-528-682-1 x A02702/rev ..

Align seg 1/1 to reverse of: A02702 from: 1 to: 795

188 GLYASNSERARGTHRIETHRGLYASPTHRCYSANGLUIGLURHGI 204
|||||
132 GGAATTCATCAAGACATCAAGAGCTGACTGTATGAGAGACCA 83
204 naslnserthrliethrglyaspthrcysanlgluiglyarg 221
|||||
82 GAATCTGACCAATATATCTCAGGGAATATCAATCAAAAGTTAAGGC 33
221 lnlepheserargthriethrglyaspthrcysanlgluigluval 230
|||||
32 AGATATTTTCAGACTATCAGCTCAAGSTN 4

seq_name: gb_pat:A06082

seq_documentation_block:

LOCUS A06082 791 bp DNA linear PAT 25-MAY-1993


```

alignment_block:
US-09-528-682-1 x A06083/rev ...
Align seg 1/1 to reverse of: A06083 from: 1 to: 791

188 gGyAsuSerSerArgThrIleThrGlyAspTrpCysAsnGluGluThrG1 204
||||| :|||: |||||
128 GGAATTCATCATCAGACAAATCAGACGCTCTACTGTGAATAGAGAGACCA 79
|||||
204 nasnleuSerThrIleTyrlLeuArgGluTyrlGlnSerLysValLysArg 221
|||||
78 GAATCGACACCAATATATCTCAGGGAATATCAATCAAAAGTTAAGAGGC 29
|||||
221 InlleheSerAspTrpGln 227
|||||
28 AGATATTTTCAGACTATWCAG 9

seq_name: gb_ba:VIBTOX

seq_documentation_block:
LOCUS VIBTOX 512 bp DNA linear BCT 26-APR-1993
DEFINITION Vibrio cholerae toxA and toxB genes for cholera enterotoxin
subunits A2 (gamma) and B.
ACCESSION K01170.1 GI:155296
VERSION K01170
KEYWORDS enterotoxin.
SOURCE Vibrio cholerae (El Tor biotype strain 62746) DNA, clone pCVD002.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 512)
AUTHORS Lockman,H. and Kaper,J.B.
TITLE Nucleotide sequence analysis of the A2 and B subunits of Vibrio
cholerae enterotoxin
J. Biol. Chem. 258, 13722-13726 (1983)
84061784
COMMENT The termination codon of subunit A2 overlaps the initiation codon
of subunit B by 4 bp. The correct translation of the 2 subunits is
the result of a shift in reading frames at this junction. Previous
protein structural analysis has shown that the A2 subunit is bound
to a precursor of the A subunit. Therefore, the 'met' at position 1
is not an initiator.
FEATURES
source Location/Qualifiers
1..512
/organism="Vibrio cholerae"
/db_xref="taxon:566"
1..141
/note="enterotoxin subunit A2"
/codon_start=1
/transl_table=1
/protein_id="AA27572.1"
/db_xref="GI:155297"
/transl_table="MSNTCDEKTSLSGVKFLDEYQSKVKRQIFSGYOSIDITHNRIRD
EL"
138..512
/note="enterotoxin subunit B"
/codon_start=1
/transl_table=1
/protein_id="AA27573.1"
/db_xref="GI:155298"
/transl_table="MIKIKGVFPTVLSSAVAHGTPONTIDCAEYHNTQIYTLNDK
IFSTESLACKREMAITTFKNGAIFQVEVPSCHIDSQKAIERMKDTLRIAYLLEAK
VEKLCVNNKTPHAIATISMAN"
141..200
/note="enterotoxin subunit B signal peptide"
201..509
/product="enterotoxin subunit B mature peptide"
BASE COUNT 199 a 74 c 88 g 151 t
ORIGIN 188 bp upstream of NdeI site.

Alignment_scores:
Quality: 157.00 Length: 44
Ratio: 3.925 Gaps: 0

```


Percent Similarity: 90.909 Percent Identity: 65.909

Alignment block:

US-09-528-682-1 x VIBTOX ..

Align seg 1/1 to: VIBTOX from: 1 to: 512

```
197 ASPTThCvASnglUgIuThrGlnAsnLeuSerThIleTyrlEuArgI 213
      :::::::::::::::::::::::::::::::::::::::::::: ||
7 AATACCTGGATGATAAAACCAAGCTGATGATAAATTCCTTGACGA 56
      |||||||
213 uTyrgInSerLySValLySArgGlnIlePheSerAspTyrgInSerGluV 230
      |||||||
57 AATACCAATCTAAAGTTAAAGACAATATTTTCAGGCTATCAATCTGATA 106
      |||||||
230 aASpIleTyraSnaRgIleArGaSPGluLeu 240
      ::||| ::::::::::::::::::::||
107 TTGATACACATATAGATTAGATTAGATGAATTA 138
```

seq_name: gb_ba:VIBCTA2

seq_documentation_block:

LOCUS VIBCTA2 251 bp DNA linear BCT 26-APR-1993

DEFINITION Vibrio cholerae cholera toxin gene: subunit A gamma-chain 3' end 6

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.

subunit B 5' end.


```

/translation="Gt:4774311"
/feature="MOPVNMOPNYKDPVNGVDIAIYIKIPNAGOMQPVKAPKINKINW
IPERDFTTNEEGBDLNPPEAKOVPSYISTYSTIDTNEKDNTLKVTTKPERISYSD
LGRKLSTIVRGIFPWSGSTIDTELAKIDINCIVIDPDGSYSSEELNLVIIGPSADI
IOECSFGEHAEVFNENGGSTQYIFESDFTEFESESLEVDNPLLGAKRATDA
VTTLAHELHAHGHRLYGAIAMPNRFFKNATAYEMSGLSEFBLRFGGHDKAFIDS
LOJNEFLYYLNKFADIASILNKAKSVGTASIOYKRNFKEXYLISEPTSGKFSVD
KLRFDKYLMLETIEDNVKPFKKYNRTGYLFEDFAVRKINVPKVNTIYDGFML
RNNINANFNQMETETNNMATEVKIKNTKTGFEPEFKLLCVGITSIRKSISDKGNKTE
GRDDGALNDICIKVNNMDLPSPSEDNFPTDLNKGEBITSDTNIEAEBNISDLLTOO
YILTFNDNEPENISIBNLSDDIIIGOLEMPNIBEPNGKRIELDKITMYHLRAQEF
EHGSRIALTNNSVEALDNPSRYTTFPSDYKVVKNATBAAMEVLUGEOLVYDFDE
TSEVSTDKIADITILLPIYIGPALNIGNMMYKDDPVSALLFSGAVILLEIPETAIPV
LGFEFYSIANKVLYIQTIDNALSKNEMKEDEYKYIVTNMLAKVOTQIDLIRKKK
EALENOAETRAKAIINYQYTREEKENNINIYIIDSLKSIENSINKAMININKPLNOC
SVSIXMSMTIPLYGKRLREDDASIKDALIKYIYIDNRGTLIGOVDRLEKDYNNLTSDOI
PFOLSKRVYDNRRLSTFTYEIKSRPOSKVQRIFSGIOSDIDITHNRJKDEL"
BASE COUNT      974 a    418 c    509 g    823 t
ORIGIN
alignment_scores:
    Quality:     109.50          Length:       42
    Ratio:       3.318           Gaps:         1
    Percent Similarity: 78.571   Percent Identity: 54.762
alignment_block:
US-09-528-682-1 x A69697 ..
Align seg 1/1 to: A69697 from: 1 to: 2724
200 AsnglucgtThrglnAsnLeusErThrIleYrLneuArgcglutrr...cl 215
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2596 AATCAAGAATTATTAATCTACATTACTGATTAATAATCAAGCCCTCA 2645
nserlysvAllysArvgcInllepheserapygrInsrgluvalaspr 232
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2646 ACTCTAAGTTAAAGCACAATATTTTCACGCCTATCATCTGATTAATGA 2695
232 lETrYsnArqgIlleArqsApglueu 240
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2696 CACATTAATGAATTAAGATGAATTA 2721
seq_name: gb_pat:BD009884
seq_documentation_block:
LOCUS      BD009884                2724 bp            DNA             linear      PAT 31-JAN-2002
DEFINITION Recombinant toxin fragments.
ACCESSION  BD009884
VERSION    BD009884.1 GI:18638257
KEYWORDS   JP 2001502890-A/8.
SOURCE     unidentified.
ORGANISM   unidentified.
REFERENCE  1 (bases 1 to 2724)
AUTHORS   Shone,C.C., Quinn,C.P. and Foster,K.A.
TITLE      Recombinant toxin fragments.
JOURNAL    Patent: JP 2001502890-A 8 06-MAR-2001;
MICROBIOLOGICAL RESEARCH AUTHORITY CAMR, THE SPEYWOOD LABORATORY LTD
COMMENT    OS Unidentified
PN         JP 2001502890-A/8
PD         06-MAR-2001
PE         22-AUG-1997 JP 1998510524
PR         23-AUG-1996 GB 9617671.4 13-DEC-1996 GB 9625966.5 PT
CLIFFORD CHARLES SHONE,CONRAD PADRAIG QUINN,KEITH ALAN FOSTER PC
CI2NI5/31,CI2NI1/21,CI2P21/02,C07K14/33,A61K38/16,A61K39/08 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 1..2724.
        Location/Qualifiers
1..2724 /organism="unidentified"

```

```

BASE COUNT      974 a 418 c 509 g 823 t      /db_xref="taxon:32644"
ORIGIN

alignment_scores:
  Quality: 109.50      Length: 42
  Ratio: 3.318         Gaps: 1
  Percent Similarity: 78.571      Percent Identity: 54.762

alignment_block:
  US-09-528-682-1 x BD009884      ..

Align seg 1/1 to: BD009884 from: 1 to: 2724

200 AsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr...GI 215
|||||.....:::|||||.....:::|||||
2596 AATCAAGATGTTATCTACATTTACTGATATATTAAGTCAAGGCTCA 2645

215 nserIysValIysArgGlnIlePheSerAspTyrGlnSerGluValAspI 232
2646 ACTTAAGATTAAAGACAAATATTTTCAGCGATCAATCTGATATGTGATA 2695

232 IeTyrAsnArgIleArgAspGluLeu 240
:::|||||.....:::|||||
2696 CACATTAATAGAAATTAAGATGAATTA 2721

seq_name: gb_ba:VIBCTA1

seq_documentation_block:
LOCUS      VIBCTA1      255 bp      DNA      linear      BCT 26-APR-1993
DEFINITION Vibrio cholerae cholera gene: subunit A alpha-chain 5' end.
ACCESSION J01846
VERSION J01846.1 GI:155156
KEYWORDS cholera toxin.
SEGMENT 1 of 2
SOURCE Vibrio cholerae (strain El Tor 1621) DNA.
ORGANISM Vibrio cholerae
REFERENCE 1 (bases 1 to 255)
AUTHORS Gemmaro,M.L. and Greenaway,P.J.
TITLE Nucleotide sequences within the cholera toxin operon
JOURNAL Nucleic Acids Res. 11, 3855-3861 (1983)
MEDLINE 83246519
COMMENT clean copy provided by authors.
FEATURES
  source
    location/Qualifiers
      1..255
        /organism="Vibrio cholerae"
        /db_xref="taxon:566"
BASE COUNT      85 a 23 c 38 g 109 t
ORIGIN

alignment_scores:
  Quality: 108.00      Length: 23
  Ratio: 4.909         Gaps: 0
  Percent Similarity: 95.652      Percent Identity: 86.957

alignment_block:
  US-09-528-682-1 x VIBCTA1      ..

Align seg 1/1 to: VIBCTA1 from: 1 to: 255

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProCaspGluIleTyr 17
|||||.....:::|||||.....:::|||||
187 AATGATGATAGATTATATCGGCGAGATCTAGACCTCTGATGATAAATAA 236

17 sarSerGlyGlyLeuMet 23
|:::|||||.....:::|||||
237 GCAGTCAGGTGCTTATG 255

seq_name: gb_pat:103425

```



```

seq_documentation_block:
LOCUS      103425          72 bp ss-DNA          linear      PAT 21-MAY-1993
DEFINITION Sequence 1 from Patent US 4882278.
ACCESSION  103425
VERSION    103425.1  GI:270622
KEYWORDS
SOURCE
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 72)
AUTHORS    Mekalanos,J.J.
TITLE       Non-toxicogenic Vibrio cholerae mutants
JOURNAL     Patent: US 4882278-A 1 21-NOV-1989;
            President and Fellows of Harvard College;
            Cambridge, MA
FEATURES
  source             1..72
                     /organism="unknown"
BASE COUNT          17 a      14 c      21 g      20 t
ORIGIN
alignment_scores:
  Quality: 106.00      Length: 23
  Ratio: 5.048         Gaps: 0
  Percent Similarity: 91.304  Percent Identity: 78.261
alignment_block:
US-09-528-682-1 x 103425  ..
Align seg 1/1 to: 103425 from: 1 to: 72
152 Asnleuasnllaalpaaaglaaspcltytrargleualaglphepr 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 AACTTACATATTCCTCCAGCAGCATGTATGATGACGAGCTTCCC 53
168 opRoasPhisGlnAlaIaTrp 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
54 TCCGGACCATGACGCTTGG 72
seq_name: gb_hlg:AC107423
seq_documentation_block:
LOCUS      AC107423          167751 bp      DNA          linear      HTG 13-FEB-2002
DEFINITION Homo sapiens chromosome 3q clone RP11-372M20, WORKING DRAFT
SEQUENCE    8 unordered pieces.
ACCESSION  AC107423
VERSION     AC107423.4  GI:18653537
KEYWORDS    HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 167751)
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbista,U., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,U., Bowie,S., Bivleva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Dreper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,J.H., Guereza,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs1,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,U., Li,Z., Licharge,O., Lieu,C., Liu,C., Liu,W., Ma,J.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mauboney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,
Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 167751)
Morley,K.C.
Direct Submission
Submitted (20-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 13, 2002 this sequence version replaced g1:18581385.
----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: HDRX
Center clone name: RP11-372M20
----- Summary Statistics -----
Sequencing vector: Plasmid; M7789
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 163427 bases at least Q40
Consensus quality: 165469 bases at least Q40
Consensus quality: 166917 bases at least Q20
Estimated insert size: 167289; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 5x in Q20 bases; sum-of-coverage estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
42446: contig of 42446 bp in length
42447 42546: gap of unknown length
42547 71415: contig of 28869 bp in length
71416 71515: gap of unknown length
71516 99155: contig of 27640 bp in length
99156 99255: gap of unknown length
99256 124809: contig of 25554 bp in length
124810 124909: gap of unknown length
124910 147053: contig of 22144 bp in length
147054 147153: gap of unknown length
147154 154262: contig of 7109 bp in length
154263 154362: gap of unknown length
154363 161157: contig of 6795 bp in length
161158 161257: gap of unknown length

```


* 161258 167751: contig of 6494 bp in length.
FEATURES
Source Location/Qualifiers
1.167751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3q"
/clone="RP11-372M20"
BASE COUNT 54466 a 30864 c 31135 g 50585 t 701 others
ORIGIN

alignment_scores:
Quality: 101.00 Length: 149
Ratio: 1.262 Gaps: 7
Percent Similarity: 53.691 Percent Identity: 24.832

alignment_block:
US-09-528-682-1 x AC107423 ..

Align seg 1/1 to: AC107423 from: 1 to: 167751

100 ValLeuGIValTYrSerProHISProTYrGIuGluGluValSerAlaLe 116
|||||::: ||| |||:::|
45208 GTGCTGGGTACACTGGATATCCATATGACAGACAAATGAAACT..... 45249
116 uGIyGIyIleProTYrSerGlnIleTYrGIyTYrTYrArgValAspHeg 133
45250AGAGTATCCTTCG 45262

133 IyValIleAspGIu.....ArgLeuHISArgAsnArgGIuTYrArgAsp 147
45263 CTACATACAAAGCAATCAAAATGATTAATACTTAATTTAAGAGT 45312
148 ArgTYrTYrArgAsnLeuAsnIleAlaProAlaGIuAspGIyTYrArgLe 164
45313 TCCTATTATTAACTACTAGTAAACATGGGAATACTC..... 45354
164 uAlaGIyPheProAspHisGlnAlaTrpArgGIuGluProTrpIleH 181
45355CCAGGACAT.....TGGAGGGGCAAAATTCCTTG 45385

181 IShISAlaProGInGIyCySGIyAsnSerArgTrIleThrGIyAsp 197
45386 ATCAATACCCACAGCACAGCAACCAAGCAAA.....ACTGGACAA 45429
198 ThrCysAsnGIuGluThrGlnAsnLeuSerThrIleTYrLeuArgGIuTy 214
45430 ACAGGATCATCATCAAGTTAAAT.....CATCTGCACACAA 45467
214 rGInSerLySValIySArgIln.....IlePheSerA 225
45468 AGGAACAAAGTAGACAGACACACCAAGCAAGGAAAGTATTTCACAA 45517
225 sPTyrgInSerGIuValAspIleTYrAsnArgIleArGAspGIuLeu 240
45518 ACTACTCATCTCAAGAAATTAATAATAGACTATATAAGAGACTC 45564

/SID5/gcgata/genseq/genseqn-emb1/NA193.DAT.AAO51327+	1081.00	2072.48	4.3e-107
/SID5/gcgata/genseq/genseqn-emb1/NA193.DAT.AAO51331+	1081.00	2072.48	4.3e-107
/SID5/gcgata/genseq/genseqn-emb1/NA193.DAT.AAO51330+	1080.00	2070.55	5.5e-107
/SID5/gcgata/genseq/genseqn-emb1/NA200.DAT.AAO51160+	1075.00	2060.17	2.1e-106
/SID5/gcgata/genseq/genseqn-emb1/NA1992.DAT.AAO51123+	1045.00	2002.17	3.6e-103
/SID5/gcgata/genseq/genseqn-emb1/NA2002.DAT.ABA92664+	929.00	1754.55	2.2e-89
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF74430+	910.00	1744.55	8.0e-89
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF74431+	562.00	1076.98	1.2e-51
/SID5/gcgata/genseq/genseqn-emb1/NA1995.DAT.AAO70348+	354.00	666.74	8.6e-29
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAO73460+	252.00	458.13	3.6e-17
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAO73245+	245.00	442.93	4.2e-16
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAO7265+	241.50	438.48	2.7e-16
/SID5/gcgata/genseq/genseqn-emb1/NA1999.DAT.AAZ20283+	240.00	439.58	3.3e-16
/SID5/gcgata/genseq/genseqn-emb1/NA1993.DAT.AAN30089+	166.00	300.18	1.5e-08
/SID5/gcgata/genseq/genseqn-emb1/NA1992.DAT.AAO20196-	160.50	270.15	1.1e-06
/SID5/gcgata/genseq/genseqn-emb1/NA1995.DAT.AAO68906+	160.50	270.15	1.1e-06
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF876106+	158.00	289.77	8.6e-08
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF76104+	157.00	303.51	1.5e-08
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAV62684+	109.50	181.37	0.0937
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAB66961+	99.50	169.28	0.4417
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAA51232+	99.50	168.43	0.4929
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAA589103+	94.00	154.11	3.0929
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF82079+	94.00	147.55	7.17
/SID5/gcgata/genseq/genseqn-emb1/NA1997.DAT.AAT55019+	93.00	155.60	2.56
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF82080+	93.00	146.30	8.42
/SID5/gcgata/genseq/genseqn-emb1/NA1989.DAT.AAN50171+	93.00	145.03	9.91
/SID5/gcgata/genseq/genseqn-emb1/NA1989.DAT.AAN50167+	93.00	145.03	9.91
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAA40756+	91.50	147.73	7.01
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAA49275+	91.00	155.18	2.69
/SID5/gcgata/genseq/genseqn-emb1/NA1992.DAT.AAO31271+	90.00	182.12	0.0851
/SID5/gcgata/genseq/genseqn-emb1/NA1997.DAT.AAT55018+	89.00	149.21	5.79
/SID5/gcgata/genseq/genseqn-emb1/NA1997.DAT.AAT55020+	89.00	147.87	6.19
/SID5/gcgata/genseq/genseqn-emb1/NA1996.DAT.AAN01029+	89.00	137.30	26.71
/SID5/gcgata/genseq/genseqn-emb1/NA1996.DAT.AAT515994+	89.00	137.29	26.74
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAT5199662-	89.00	66.03	2.4e+05
/SID5/gcgata/genseq/genseqn-emb1/NA1992.DAT.AAO31273+	88.00	178.25	0.1398
/SID5/gcgata/genseq/genseqn-emb1/NA1992.DAT.AAO35994+	88.00	178.25	0.1398
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAZ36726+	88.00	150.25	5.07
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAA15268+	88.00	149.35	10.81
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAA160695+	88.00	139.68	19.68
/SID5/gcgata/genseq/genseqn-emb1/NA1999.DAT.AAZ06826+	88.00	126.79	102.82
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAF61461+	86.50	106.92	1.3e+03
/SID5/gcgata/genseq/genseqn-emb1/NA2000.DAT.AAF61510+	86.50	87.17	1.7e+04
/SID5/gcgata/genseq/genseqn-emb1/NA1998.DAT.AAV33063+	86.00	178.00	0.1444
/SID5/gcgata/genseq/genseqn-emb1/NA1999.DAT.AAZ27859+	86.00	178.00	0.1444
/SID5/gcgata/genseq/genseqn-emb1/NA1999.DAT.AAZ27493+	86.00	178.00	0.1444
/SID5/gcgata/genseq/genseqn-emb1/NA1999.DAT.AAZ27495+	86.00	178.00	0.1444
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAH66219+	85.00	139.40	20.39
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAH68527+	85.00	84.27	2.4e+04
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAA65132-	84.00	125.70	118.29
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAB115413+	83.50	135.14	35.21
/SID5/gcgata/genseq/genseqn-emb1/NA1994.DAT.AAO77665+	83.50	133.98	40.87
/SID5/gcgata/genseq/genseqn-emb1/NA1991.DAT.AAO41875+	83.50	126.85	101.99
/SID5/gcgata/genseq/genseqn-emb1/NA1993.DAT.AAO84725+	83.50	126.85	101.99
/SID5/gcgata/genseq/genseqn-emb1/NA1995.DAT.AAO87268+	83.50	126.85	101.99
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAF51971+	83.00	137.30	26.69
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAB51616+	83.00	136.63	29.09
/SID5/gcgata/genseq/genseqn-emb1/NA2001.DAT.AAB636708+	83.00	136.63	29.09
/SID			

/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH104418 +	83.00	133.60	42.93	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH65686 +	77.00	133.39	44.12	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH42258 +	83.00	128.19	85.89	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH12965 +	77.00	133.29	44.66	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH67017 +	82.50	137.32	26.63	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABA05183 +	77.00	131.57	55.66	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH15399 +	82.50	127.98	88.23	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH77110 -	77.00	126.99	100.24	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAH86486 +	82.50	127.44	98.36	32	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH77109 -	77.00	126.96	100.57	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH15398 +	82.50	115.21	453.65	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH86479 +	77.00	125.97	114.22	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH16028 +	82.50	108.41	1.1e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH24907 +	77.00	125.97	114.22	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABH16028 +	82.50	108.41	1.1e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH24907 +	77.00	125.97	114.22	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH41226 +	82.50	79.44	4.4e+04	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH31863 +	77.00	120.92	218.14	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH05185 +	82.00	141.12	16.37	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH31863 +	77.00	120.92	218.14	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH19869 +	82.00	133.78	41.95	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH68353 +	77.00	114.54	494.49	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH19868 +	82.00	124.78	133.11	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH68354 +	77.00	112.38	652.65	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH19968 +	82.00	119.26	270.05	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH08234 +	77.00	108.67	1.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH10190 +	82.00	119.25	270.43	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH1477 -	77.00	95.14	6.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAH22900 +	81.50	132.54	49.18	16	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH1477 -	77.00	95.14	6.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH22900 +	81.50	112.54	639.49	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH31863 +	77.00	120.92	218.14	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH17142 +	81.00	168.33	0.4988	45	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH76599 +	76.50	118.98	1.7e+05	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAH33005 +	81.00	140.82	17.01	66	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAH44668 +	76.50	116.60	379.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH05189 +	81.00	138.91	21.72	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAH44668 +	76.50	116.60	379.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH15890 +	81.00	128.25	85.12	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAH44668 +	76.50	116.60	379.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH05189 +	81.00	128.25	85.12	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAH44668 +	76.50	116.60	379.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH54174 +	81.00	124.68	134.80	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH44132 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH540217 +	80.50	132.00	52.70	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH44132 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAH41895 +	80.50	131.60	55.44	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH504223 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH501503 +	80.00	164.92	0.7729	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH504223 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH501504 +	80.00	164.93	0.7924	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH89102 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAH36375 +	80.00	164.53	0.8122	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH502329 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH86687 +	80.00	139.14	21.10	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH17390 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH05187 +	80.00	137.37	26.46	76	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH51814 +	76.00	127.19	97.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH86478 +	80.00	134.63	37.62	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAH79491 -	76.00	127.19	97.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH24906 +	80.00	134.63	37.62	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH79491 -	76.00	127.19	97.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH29871 +	80.00	128.70	80.51	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH74927 +	76.00	124.92	130.13	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH29874 +	80.00	128.33	86.61	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAH92734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH28209 +	76.00	116.11	404.38	1
/SIDS/gcgdata/gen									

/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT07595 +	72.00	106.22	1.4e+03	3	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAS03789 +	71.00	110.55	825.26	3
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAD33738 +	72.00	102.92	2.2e+03	48	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO25783 +	71.00	109.13	870.74	4
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAD33740 +	72.00	102.76	2.2e+03	49	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAS68217 +	71.00	109.30	968.69	5
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT10079 +	72.00	102.65	2.2e+03	5	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAS89662 +	71.00	109.30	968.69	6
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH74865 +	72.00	102.11	2.4e+03	5	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAV01387 +	71.00	109.17	984.57	7
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT17184 +	72.00	101.92	2.5e+03	5	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAS94048 +	71.00	108.96	1.0e+03	8
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ18356 +	72.00	101.84	2.5e+03	54	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9681 +	71.00	108.75	1.0e+03	9
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB0254 +	72.00	101.84	2.5e+03	54	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAS59554 +	71.00	108.34	1.1e+03	10
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ18357 +	72.00	101.68	2.6e+03	55	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAO38270 +	71.00	107.08	1.3e+03	11
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ18357 +	72.00	101.68	2.6e+03	55	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAO38270 +	71.00	106.78	1.3e+03	12
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ18358 +	72.00	101.39	2.7e+03	56	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAO44407 +	71.00	106.16	1.4e+03	13
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB0256 +	72.00	101.39	2.7e+03	56	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAH84453 +	71.00	106.16	1.4e+03	14
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAD33741 +	72.00	100.29	3.1e+03	63	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAH84379 +	71.00	106.16	1.4e+03	15
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAD33743 +	72.00	98.65	3.8e+03	74	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAO31925 +	71.00	106.16	1.4e+03	16
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS5446 +	72.00	98.22	4.0e+03	7	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9663 +	71.00	106.16	1.4e+03	17
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAD33742 +	72.00	97.80	4.2e+03	80	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS28711 +	71.00	105.97	1.5e+03	18
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAAS2274 +	72.00	93.02	7.8e+03	12	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAS00690 +	71.00	105.92	1.5e+03	19
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ22189 +	72.00	92.72	8.1e+03	13	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB28710 +	71.00	105.32	1.6e+03	20
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAB08334 +	71.50	117.72	329.02	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAS89170 +	71.00	104.83	1.7e+03	21
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAS2470 +	71.50	115.22	453.57	13	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAN1089 +	71.00	104.89	1.7e+03	22
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAH98632 +	71.50	113.94	534.02	15	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAT0192 +	71.00	103.90	1.9e+03	23
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC4475 +	71.50	113.18	588.86	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS93190 +	71.00	103.58	2.0e+03	24
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAC84220 +	71.50	112.69	627.08	17	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAAT31033 +	71.00	103.35	2.1e+03	25
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAH1835 +	71.50	112.58	635.06	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI60060 +	71.00	102.13	2.4e+03	26
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS92316 +	71.50	110.03	882.32	22	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ84001 +	71.00	101.27	2.7e+03	27
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:AAAN0424 +	71.50	109.85	902.75	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB5948 +	71.00	99.56	3.4e+03	28
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS89075 +	71.50	109.24	976.44	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC6439 +	71.00	98.95	3.7e+03	29
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT12773 +	71.50	108.93	1.0e+03	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS86546 +	71.00	98.46	3.9e+03	30
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAB89120 +	71.50	108.08	1.1e+03	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABLT17100 +	71.00	97.92	4.2e+03	31
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT11015 +	71.50	107.64	1.2e+03	28	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK8349 +	71.00	96.39	5.1e+03	32
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:AAAZ0016 +	71.50	107.09	1.3e+03	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAAS9321 +	71.00	95.93	5.4e+03	33
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS8196 +	71.50	107.09	1.3e+03	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAS8571 +	71.00	95.93	5.4e+03	34
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS9038 +	71.50	106.51	1.4e+03	31	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO38218 +	71.00	94.07	6.8e+03	35
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9436 +	71.50	105.84	1.5e+03	3	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAS7202 +	71.00	93.75	7.1e+03	36
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAAS2340 +	71.50	105.50	1.6e+03	34	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9038 +	71.00	93.69	7.2e+03	37
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT12520 +	71.50	105.25	1.6e+03	3	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAD0472 +	71.00	93.69	7.2e+03	38
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT13149 +	71.50	103.13	2.1e+03	4	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAT7335 +	71.00	93.12	7.7e+03	39
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAB09494 +	71.50	103.11	2.1e+03	4	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS0387 +	71.00	93.12	7.7e+03	40
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT12772 +	71.50	102.24	2.4e+03	4	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9620 +	71.00	86.14	1.9e+04	41
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT11014 +	71.50	102.12	2.4e+03	4	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH27118 +	71.00	74.19	8.6e+04	42
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT11220 +	71.50	101.13	2.8e+03	5	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF22283 +	71.00	71.91	1.2e+05	43
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAB4352 +	71.50	98.75	3.7e+03	6	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF22283 +	71.00	71.91	1.2e+05	44
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS59556 +	71.50	98.44	3.9e+03	6	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAH33476 +	71.00	92.04	8.9e+03	45
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT13148 +	71.50	98.28	4.0e+03	6	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC64507 +	71.00	92.03	8.9e+03	46
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT17119 +	71.50	92.36	8.5e+03	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC64507 +	71.00	92.03	8.9e+03	47
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABLT18074 +	71.50	92.36	8.5e+03	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAT7335 +	71.00	91.83	9.1e+03	48
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH89112 +	71.50	83.70	2.6e+04	2	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9038 +	71.00	91.83	9.1e+03	49
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF11052 +	71.50	67.34	2.0e+05	14	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAD23451 +	71.00	91.51	9.5e+03	50
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF12752 +	71.50	67.34	2.0e+05	14	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAD7358 +	71.00	90.26	1.1e+04	51
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB4353 +	71.50	67.34	2.0e+05	14	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS0380 +	71.00	90.26	1.1e+04	52
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS15138 +	71.50	66.90	2.2e+05	14	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS9620 +	71.00	86.14	1.9e+04	53
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS15151 +	71.50	66.90	2.2e+05	14	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLT17118 +	71.00	74.19	8.6e+04	54
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF21273 +	71.50	66.67	2.2e+05	15	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF22283 +	71.00	71.91	1.2e+05	55
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAAB03041 +	71.50	36.35	3.3e+06	29	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAF22283 +	71.00	71.91	1.2e+05	56
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS11233 +	71.00	150.22	5.10	40	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH07082 +	70.50	119.05	277.32	57
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS11238 +	71.00	150.22	5.10	40	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH07082 +	70.50	115.31	482.88	58
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS11144 +	71.00	150.22	5.10	40	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS9451 +	70.50	114.73	442.84	59
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS51144 +	71.00	120.80	5.10	40	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC9273 +	70.50	114.44	504.80	60
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS52738 +	71.00	120.80	5.10	40	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC9273 +	70.50	114.38	504.80	61
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC5240 +	71.00	119.95	247.11	76	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC9273 +	70.50	114.37	505.35	62
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH76938 +	71.00	119.78	252.63	7	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAO2011 +	70.50	112.83	615.64	63
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAH81817 +	71.00	118.87	283.71	85	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC9788 +	70.50	112.70	626.66	64
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAH1928 +	71.00	117.64	332.34	96	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAO82821 +	70.50	112.64	631.32	65
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAH21396 +	71.00	117.09	356.63	10	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS3393 +	70.50	111.32	747.22	66
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS33961 +	71.00	116.76	371.96	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH06177 +	70.50	111.00	778.89	67
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC66450 +	71.00	115.15	426.72	12	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB62001 +	70.50	110.56	824.51	68
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAF60370 +	71.00	114.30	510.39	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS51643 +	70.50	109.31	966.89	69
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH01911 +	71.00	113.33	577.88	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB64333 +	70.50	109.22	978.51	70
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH80247 +	71.00	112.90	610.77	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS54405 +	70.50	108.50	109.22	71
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS92463 +	71.00	112.90	610.77	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAC99121 +	70.50	108.50	1.1e+03	72
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS92467 +	71.00	112.90	610.77	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAAB6164 +	70.50	108.45	1.1e+03	73
/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAAS54376 +	71.00	111.01	778.14	1	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAO05917 +	70.50	108.09	1.1e+03	74

/SID5/gcgdata/geneseq/geneseqn-emb1/NA197.DAT:AA74269 +	70.50	107.94	1.2e+03	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA51819 +	70.00	104.26	1.8e+03	22
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA77052 +	70.50	107.94	1.2e+03	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA54951 +	70.00	104.02	1.9e+03	22
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA23925 +	70.50	106.44	1.4e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA25786 +	70.00	103.80	2.0e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB119383 +	70.50	106.43	1.4e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB115440 +	70.00	103.62	2.0e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA62011 +	70.50	106.36	1.4e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA091362 +	70.00	103.51	2.0e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA15824 +	70.50	106.32	1.4e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA101304 +	70.00	103.29	2.1e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA455835 +	70.50	105.76	1.5e+03	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA57377 +	70.00	102.92	2.2e+03	22
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB106777 +	70.50	105.70	1.5e+03	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA499908 +	70.00	102.86	2.2e+03	22
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA48349 +	70.50	105.12	1.7e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB115034 +	70.00	102.76	2.2e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB127356 -	70.50	103.63	2.0e+03	34	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB126659 +	70.00	102.46	2.5e+03	34
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ48497 +	70.50	103.51	2.0e+03	34	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA570241 +	70.00	102.01	2.5e+03	34
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA793651 +	70.50	103.47	2.0e+03	34	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA573824 +	70.00	102.01	2.5e+03	34
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB127205 +	70.50	103.41	2.0e+03	34	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA574283 +	70.00	102.01	2.5e+03	34
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA522678 -	70.50	102.89	2.2e+03	38	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA574893 +	70.00	102.01	2.5e+03	38
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA434193 +	70.50	102.50	2.3e+03	40	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA578725 +	70.00	102.01	2.5e+03	40
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA43498 +	70.50	102.03	2.5e+03	40	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA578940 +	70.00	102.01	2.5e+03	40
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA158274 +	70.50	101.46	2.6e+03	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA581661 +	70.00	102.01	2.5e+03	44
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA570098 +	70.50	100.73	2.9e+03	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA584092 +	70.00	102.01	2.5e+03	44
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA481541 -	70.50	100.67	2.9e+03	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV3697 +	70.00	101.97	2.5e+03	44
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB119382 -	70.50	100.34	3.1e+03	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB114101 +	70.00	101.65	2.6e+03	44
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA82427 +	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ14530 +	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ78074 +	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ63617 +	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ93651 +	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA574485 +	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA522423 +	70.50	99.09	4.0e+03	57	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA574307 +	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA220606 +	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA578719 +	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA114547 +	70.50	95.92	5.4e+03	72	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA578953 +	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB104502 +	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV62177 +	70.00	101.04	2.8e+03	72
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA33437 +	70.50	95.54	5.7e+03	72	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ35118 +	70.00	99.92	3.2e+03	72
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AB127204 +	70.50	94.38	6.6e+03	8	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV74545 +	70.00	99.40	3.4e+03	8
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ81514 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB119800 +	70.00	94.26	6.7e+03	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA730787 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB129748 +	70.00	94.10	6.7e+03	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV60125 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB120244 +	70.00	92.26	8.6e+03	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV42684 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AB133689 +	70.00	90.68	1.1e+04	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAZ92765 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA527648 -	70.00	87.66	1.6e+04	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ92785 +	70.50	87.45	1.6e+04	16	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AB19256 +	70.00	87.51	1.6e+04	16
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ92892 +	70.50	85.89	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB109378 -	70.00	86.27	1.9e+04	1
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA54561 +	70.50	84.95	2.3e+04	21	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV52233 +	70.00	79.18	4.6e+04	21
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA513031 +	70.50	84.73	2.3e+04	21	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV78804 +	70.00	78.84	4.8e+04	21
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV51569 +	70.50	80.51	3.9e+04	32	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV21610 +	70.00	55.28	8.6e+05	32
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV52204 +	70.50	79.81	4.2e+04	35	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV30459 +	70.00	50.94	1.4e+06	35
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA511143 +	70.00	148.28	6.53	40	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA533208 +	69.50	111.66	7.15	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA511143 +	70.00	148.28	6.53	40	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA532209 +	69.50	111.66	7.15	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV90061 +	70.00	123.48	157.11	45	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA577369 -	69.50	112.97	605.40	45
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAZ10824 +	70.00	122.86	172.42	48	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA531321 +	69.50	112.09	677.37	48
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA521980 +	70.00	119.84	250.61	64	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA531427 +	69.50	111.66	7.15	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA52436 +	70.00	118.96	280.78	7	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA533208 +	69.50	111.66	7.15	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA710681 +	70.00	118.65	292.97	72	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA532209 +	69.50	111.66	7.15	51
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA710682 +	70.00	117.63	332.99	77	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA56150 +	69.50	110.78	801.45	77
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH29524 -	70.00	115.28	449.86	10	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH15771 +	69.50	110.78	801.45	10
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAQ24847 +	70.00	114.57	492.81	10	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH14030 +	69.50	110.65	811.87	10
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ24679 +	70.00	113.90	536.68	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH14030 +	69.50	110.65	811.87	11
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ86890 +	70.00	112.14	673.39	13	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH75173 +	69.50	110.54	826.10	13
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ83252 +	70.00	111.88	732.46	14	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH77715 +	69.50	109.97	888.88	14
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ51620 +	70.00	110.69	810.31	15	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ53394 +	69.50	109.96	957.40	15
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ51620 +	70.00	110.69	810.31	15	/SID5/gcgdata/geneseq/geneseqn-emb1/NA198.DAT:AAQ52478 +	69.50	108.75	999.56	15
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA50585 +	70.00	110.67	813.02	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH44842 -	69.50	108.61	1.1e+03	1
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA569603 +	70.00	108.62	1.1e+03	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA577427 -	69.50	108.61	1.1e+03	1
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA580196 +	70.00	108.60	1.1e+03	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA577573 +	69.50	108.61	1.1e+03	1
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ20462 +	70.00	107.93	1.2e+03	20	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA571119 -	69.50	108.61	1.1e+03	20
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH20438 +	70.00	106.73	1.3e+03	23	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB106815 +	69.50	108.38	1.1e+03	23
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA747072 +	70.00	106.58	1.4e+03	23	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA588910 -	69.50	108.05	1.1e+03	23
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AA59989 +	70.00	105.93	1.5e+03	24	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA52442 +	69.50	108.05	1.1e+03	24
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ80750 +	70.00	105.93	1.5e+03	24	/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ53305 +	69.50	107.63	1.2e+03	24
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAV65780 +	70.00	105.93	1.5e+03	24	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AA510662 +	69.50	107.63	1.2e+03	24
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ10163 +	70.00	105.18	1.6e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB104519 +	69.50	106.15	1.5e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA199.DAT:AAQ01792 +	70.00	104.66	1.8e+03	26	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV79272 +	69.50	105.12	1.7e+03	26
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB116672 +	70.00	104.61	1.8e+03	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA584427 +	69.50	105.12	1.7e+03	2
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AB110571 +	70.00	104.61	1.8e+03	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV9945 +	69.50	105.00	1.7e+03	2

/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB64212 +	69.50	103.85	1.9e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA89141 +	69.00	73.65	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB106517 +	69.50	101.75	2.6e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA81501 -	69.00	73.64	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.AAO27486 -	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA889142 +	69.00	73.63	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO33879 -	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA85467 +	69.00	73.28	9.7e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT92868 +	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA58471 +	69.00	71.61	1.2e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT92701 +	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI19214 +	69.00	70.88	1.3e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAX91114 +	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF54868 +	69.00	66.32	1.8e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAQ12520 +	69.50	97.27	4.4e+03	51	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF22288 +	69.00	66.97	2.1e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.AAO25446 -	69.50	97.27	4.5e+03	52	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF21611 +	69.00	53.35	1.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB16638 -	69.50	96.50	5.0e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA68533 -	69.00	53.35	1.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT75715 -	69.50	95.91	5.4e+03	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA81489 +	69.00	44.41	2.4e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABI33333 +	69.50	92.99	7.8e+03	80	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA57637 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAL03113 +	69.50	89.49	1.2e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA27059 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAV62160 -	69.50	89.13	1.3e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA05691 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK73194 -	69.50	88.83	1.3e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA31306 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS96639 -	69.50	88.79	1.3e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI15627 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAS66133 -	69.50	88.29	1.4e+04	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI37203 +	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59599 +	69.50	87.18	1.7e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX61794 +	68.50	118.04	315.90	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59599 +	69.50	82.71	2.9e+04	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI15649 +	68.50	116.86	367.20	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB120416 -	69.50	79.46	4.4e+04	30	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX61793 +	68.50	116.12	404.17	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAA81520 +	69.50	78.07	5.3e+04	34	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAB24215 +	68.50	114.42	502.15	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO54386 +	69.50	74.91	7.9e+04	46	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB24015 +	68.50	113.65	554.22	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA51135 -	69.50	74.67	8.1e+04	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB56142 +	68.50	110.46	835.07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF30757 +	69.50	65.52	2.6e+05	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA06734 +	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV62176 -	69.50	62.67	3.6e+05	15	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA41457 +	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD25119 -	69.50	54.31	9.6e+05	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA81721 +	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH68532 +	69.50	54.31	9.6e+05	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA23392 +	68.50	110.01	884.43	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH68532 +	69.50	28.31	3.6e+06	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF18014 +	68.50	109.72	917.71	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA51135 -	69.00	146.35	8.37	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV11921 +	68.50	109.28	970.99	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF71262 +	69.00	119.13	274.64	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA06561 +	68.50	109.17	985.13	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS73882 -	69.00	119.12	278.44	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF88330 +	68.50	109.04	1.0e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS73882 -	69.00	118.12	312.59	62	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA40837 +	68.50	108.75	1.0e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH53653 +	69.00	116.52	383.58	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI59400 +	68.50	107.27	1.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH53653 +	69.00	115.49	437.95	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX22899 +	68.50	107.20	1.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC33686 +	69.00	114.29	505.37	91	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI61186 +	68.50	107.16	1.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC33686 +	69.00	113.10	555.37	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI61187 +	68.50	107.16	1.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF67889 +	69.00	112.03	682.91	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI59401 +	68.50	106.62	1.4e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF67889 +	69.00	110.29	853.65	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT65633 +	68.50	105.80	1.6e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAX77550 +	69.00	110.29	853.65	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14479 +	68.50	104.86	1.7e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAX77550 +	69.00	109.61	930.85	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT69697 +	68.50	103.54	2.0e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF67151 +	69.00	109.38	958.93	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF51928 +	68.50	103.54	2.0e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB033716 +	69.00	109.35	962.35	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS84519 +	68.50	103.32	2.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF07094 +	69.00	108.77	1.0e+06	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA51927 +	68.50	103.14	2.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAN93090 +	69.00	108.69	1.0e+06	15	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA53098 -	68.50	102.72	2.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO10060 +	69.00	107.91	1.2e+06	17	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA51693 +	68.50	102.59	2.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI17755 +	69.00	107.54	1.2e+06	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA52677 +	68.50	102.47	2.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03129 +	69.00	107.11	1.3e+06	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14455 +	68.50	101.40	2.7e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT40701 +	69.00	106.36	1.4e+06	19	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78939 +	68.50	100.83	2.9e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS89745 +	69.00	104.65	1.8e+06	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82654 +	68.50	100.83	2.9e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV53376 +	69.00	104.57	1.8e+06	23	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82654 +	68.50	100.83	2.9e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS84101 +	69.00	104.34	1.8e+06	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS49236 +	68.50	100.16	3.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA47605 +	69.00	103.85	2.0e+06	25	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS88406 +	68.50	100.07	3.2e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI25999 +	69.00	103.82	2.0e+06	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS88406 +	68.50	99.81	3.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH16365 +	69.00	103.27	2.1e+06	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.AAN70681 +	68.50	99.68	3.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03666 +	69.00	102.35	2.4e+06	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14478 +	68.50	99.67	3.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH14636 -	69.00	101.34	2.6e+06	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS68181 +	68.50	99.35	3.5e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC86480 +	69.00	100.86	2.9e+06	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX20279 +	68.50	98.83	3.7e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS66420 +	69.00	100.25	3.1e+06	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03577 +	68.50	98.41	3.9e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI17753 +	69.00	99.78	3.3e+06	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.AAB150516 +	68.50	98.21	4.0e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI24348 -	69.00	99.16	3.6e+06	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI15693 +	68.50	98.05	4.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH54332 -	69.00	98.82	3.7e+06	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH98351 +	68.50	97.65	4.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC77490 -	69.00	98.82	3.8e+06	42	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA53019 -	68.50	97.65	4.3e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI25998 +	69.00	97.69	4.3e+06	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAA53019 -	68.50	97.37	4.5e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAA55474 +	69.00	95.19	5.9e+06	58	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABI9832 +	68.50	96.96	4.7e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV16307 +	69.00	95.19	5.9e+06	58	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AAA50995 +	68.50	95.62	5.6e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF79751 +	69.00	94.25	6.7e+06	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI08295 +	68.50	95.16	5.9e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH57445 +	69.00	94.24	6.7e+06	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV20464 +	68.50	92.45	8.4e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAZ4761 +	69.00	94.20	6.7e+06	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI08294 +	68.50	91.04	1.0e+07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV16345 +	69.00	94.15	6.8e+06	65	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT101788 +	68.50	88.78	1.3e+07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59595 -	69.00	87.46	1.6e+07	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS87216 -	68.50	87.90	1.5e+07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59595 -	69.00	85.15	2.1e+07	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03576 -	68.50	86.12	1.9e+07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59595 -	69.00	82.56	3.0e+07	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI07768 -	68.50	76.90	6.1e+07	

/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH54688 - 67.50 97.48 4.4e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAO85414 + 67.50 97.46 4.4e+03 35
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAO6246 + 67.50 97.46 4.4e+03 35
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAO3609 + 67.50 97.46 4.4e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH8969 + 67.50 97.06 4.7e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH8936 + 67.50 96.61 4.9e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLO5233 + 67.50 96.14 5.2e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABLO7507 + 67.50 96.01 5.3e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABLI5138 - 67.50 95.85 5.4e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ADD24312 + 67.50 95.82 5.5e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK52035 + 67.50 95.65 5.6e+03 4

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAH50205

seq_documentation_block:
 ID AAN50205 standard; DNA; 777 BP.

AC AAN50205;

DT 30-OCT-1991 (first entry)

DE Sequence of the pig scours heat labile toxin (LT) LTA gene.

XX Pig scours vaccine; toxin; diarrhoea; ss.

OS E.coli NCIB 11932.

XX Key Location/Qualifiers

FT CDS 1..777

FT /*tag= a

PN EP145486-A.

XX 19-JUN-1985.

XX 12-DEC-1984; 84EP-0308620.

XX 12-DEC-1983; 83GB-0033131.

XX (GLAX) GLAXO GROUP LTD.

XX PI Hayes MV, Harford S, Ross GW;

XX WPI: 1985-148358/25.

XX P-PSDB: AAP50190.

XX New toxoid as inactivated form of toxin for use in vaccines - is

XX obd. from organism transformed by gene

XX Disclosure; Fig 1; 61pp; English.

XX AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of

XX the site directed mutant SDM1 (see AAN50206) is inactive. The

XX inventors claim a vaccine prep. active against pig scours which

XX contains an inactivated LTA component, together with additional K88

XX antigens opt. with whole cells comprising the antigens or contg. the

XX inactivated LTA.

XX Sequence 777 BP; 255 A; 136 C; 164 G; 222 T; 0 other;

XX

XX

XX

XX

XX

XX

55 AATGGGACGATTATACCGTGTGACTGTAGACCCCGCATTAATAA 104
 17 sarserglyglyleumetproarglyhisangluityrphespargg 34
 105 ACCTTCGGGGGTCTTATGCCAGAGGGCATATGAGTACTTCGATAGAG 154
 34 lythrlnmetlanlleasnleutyrsphissalarglythrlnmr 50
 155 GAACCTCAATGAAATATTTATCTTATGATCAGCGAGAGCAACCAAC 204
 51 GlypheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
 205 GCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 254
 67 gserAlahisIleuAlaglylnserIleuSerGlyTyrSerThrTyrT 84
 255 AAGTGCTCACTTAGCAGACAGCTATATATATATATATATATATATAT 304
 84 yrlletYrValIleAlatThrAlaProAsnMetPheAsnValAsnspAl 100
 305 ATATATATATGTTATAGCAGACACCAATATGTTATATATATATATAT 354
 101 LeuGlyValTyrSerProHisProTyrGluGluValSerAlaLeuG 117
 355 TTAGCGGTATACAGCCCTCACCATATGAAAGAGGAGGTTTCGCTTAG 404
 117 yglylleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnphely 134
 405 TGGATATCATATATCTCATATATATGATGATGATGATGATGATGATG 454
 134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
 455 TGATGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 505 AGAAATCTGAATATAGCTCCGGCAGAGGATGCTTACAGATTAGCAG 554
 167 eProProAspHisGlnAlaTPrArgGluGluProThrIleHisAlaP 184
 555 CCCACCGATCCCAAGCTTGGAGAGAAAGAACCTGGATTCATCATGAC 604
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
 605 CACAAGGTGTGGAATTCATCAACAACATCAGATGATGATGATGATG 654
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 655 GAGGAGACCCAGAAATCTGAGCAATATATCTCAGGGATATCATCAMA 704
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 705 AGTTAAGAGCGAGATATTTTCAGACTATCAGTACAGAGGTTCATATA 754
 234 sPheAspGluLeu 240
 755 ACAGATTCGGGATGATTA 774

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAH50206

seq_documentation_block:

ID AAN50206 standard; DNA; 777 BP.

AC AAN50206;

DT 30-OCT-1991 (first entry)

DE Sequence of the pig scours heat labile toxin (LT) LTA gene of the

DE site directed mutant SDM1.

XX Pig scours vaccine; toxin; diarrhoea; ss.

OS E.coli NCIB 11932.


```

XX Key Location/Qualifiers
FH CDS 1..777
FT /*tag= a
FT misc_difference 235..237
FT /*tag= b
FT /*note= "TCC in native SQ"
XX
XX EPI45486-A.
XX
XX 19-JUN-1985.
XX
XX 12-DEC-1984; 84EP-0308620.
XX
XX 12-DEC-1983; 83GB-0033131.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Hayes MV, Harford S, Ross GW;
XX
XX WPI: 1985-148358/25.
XX P-PSDB: AAP50191.
XX
XX New toxoid as inactivated form of toxin for use in vaccines - is
XX obtd. from organism transformed by gene
XX
XX Example: Fig 2; 61pp: English.
XX
XX AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of
XX the site directed mutant SDM1 (see AAN50206) is inactive. The
XX inventors claim a vaccine prepn. active against pig scour which
XX contains an inactivated LTA component, together with additional K88
XX antigens opt. with whole cells comprising the antigens or contg. the
XX inactivated LTA.
XX
XX Sequence 777 BP; 255 A; 135 C; 164 G; 223 T; 0 other;
XX
alignment_scores:
Quality: 1298.00 Length: 240
Ratio: 5.431 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 99.583
alignment_block:
US-09-528-682-1 x AAN50206 ..
Align seg 1/1 to: AAN50206 from: 1 to: 777
1 AanglyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
|||||
55 AATGGCGACAGATTAACCGTGTGACTTACAGCCCCGATGAAATATA 104
17 SAArgSerGlyGlyLeuMetProArgGlyHisAnGluTyrPheAspArg 34
|||||
105 ACGTCCGAGAGCTTATGCCAGAGCGCATATAGTACTTCGATAGAG 154
34 LyThGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCAAAATGAATATTAATCTTATGATCAGCGAGAGAAACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
205 GCCTTGTTCAGATATGATGACGATATGTTTCACTTCCTTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
255 AAGTGTCTACTAGCAGACAGTCTATATTATACAGATATTCACACTTAC 304
84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATATGTTATAGCAGACACAAATATGTTTATATGTTAATGATGA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117

```

```

|||||
355 TTAGCGGTATACAGCCCTCACCATTATGAACAGAGAGCTTTCGCTTAG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||
405 TGGATATACCATATTTCTCAGATATATGATGATGTTCTGTAAATTTTGG 454
134 AIleAspGluArgLeuHisArgAsnArgGluTyrTyrArgSerArgTyr 150
|||||
455 TGATTGATGACGATTATCATCGTAACAGGGAATATAGACACCGGTATTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||
505 AGAAATCTGAATATAGCTCCGCGAGAGATGCTTACAGATTACAGAGTT 554
167 eProProAspHisGlnAlaTyrArgGluGluProTyrIleHisAlaP 184
|||||
555 CCACCCGATCACACAGCTTGAGAGAGAACCCCTGATTCATCATGCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
|||||
605 CACAAAGTTGTGGAAATTCATCAAGAACATCACAGGTGATACTTGTAA 654
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
655 GAGAGACCCAGAAATCTGACACAAATATATCTCAGGGAATATCAATCAA 704
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
705 AGTTAAGAGGAGATATTTTACAGACTATCATCAGCAGAGTTGACATATA 754
234 sNArgIleArgAspGluLeu 240
|||||
755 ACAGAAATTCGGGATGAATTA 774
seq_name: /SID55/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAS01505
seq_documentation_block:
ID AAS01505 standard; DNA; 1514 BP.
XX
XX AAS01505;
XX
XX 29-AUG-2001 (first entry)
XX
XX E. coli heat-labile enterotoxin (LT) mutant LTS63Y DNA.
XX
XX Heat-labile enterotoxin; LT; LTS63Y; LTdel110/112; mutant;
XX detoxified and immunologically active protein; Adp-ribosylation; Gs;
XX endotoxin; diarrhoea; ds.
XX
XX Escherichia coli strain K88ac.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH CDS 164..1311
FT /*tag= a
FT /*product= "LTS63Y mutant protein"
FT /*transl_except= (pos:938..939,aa:Met)
FT /*note= "This codon has an apparent 1 nucleotide
deletion which alters the reading frame"
XX
XX WO200119998-A1.
XX
XX 22-MAR-2001.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX

```


DR WPI: 2001-281524/29.
 DR P-PSDB: AAU00506.
 PT New detoxified mutants of *Escherichia coli* heat-labile enterotoxin
 PT useful as vaccine for preventing and treating diarrhoea, and as adjuvant
 PT for antibody production
 PS
 PS
 PS Claim 3: Page 41-42; 48pp; English.
 XX
 CC The present sequence encodes for *Escherichia coli* heat-labile
 CC enterotoxin (LT) mutant LRS63Y. LRS63Y and LTdel110/112 (AAU00507)
 CC are two novel detoxified and immunologically active proteins (LT
 CC mutants) derived by site-directed mutagenesis of the A1 subunit of wild
 CC type LT. The substitution of Ser to Tyr at position 63 in LRS63Y blocks
 CC NAD-binding. Deletion of Glu residues at positions 110 and 112 in
 CC LTdel110/112 eliminate the enzymatic activity of LT. The A1 subunit of
 CC wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
 CC regulates cAMP levels. The resulting increase in cAMP is the cause of
 CC diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
 CC of mutant heat-labile enterotoxins LRS63Y and LTdel110/112 were tested.
 CC Groups of mice were immunised with LRS63Y or LTdel110/112. The control
 CC groups received phosphate buffered saline (PBS) alone. The serum and
 CC faecal antibody titres to LT were determined. The results showed that
 CC mice immunised with LRS63Y or LTdel110/112 contained high and
 CC comparable level of anti-LT antibodies in sera and faecal extracts
 CC compared with those immunised with wild-type LT. The LT mutants are
 CC useful as a vaccine for preventing and treating diarrhoea and as an
 CC adjuvant for antibody production.
 XX
 SQ Sequence 1514 BP; 510 A; 277 C; 298 G; 429 T; 0 other:

alignment_scores:

Quality: 1298.00 Length: 240
 Ratio: 5.431 Gaps: 0
 Percent Similarity: 99.583 Percent Identity: 99.583

alignment_block:

US-09-528-682-1 x AAS01505 ..

Align seg 1/1 to: AAS01505 from: 1 to: 1514

1 AsnGLYAspArgLeuTyrArgAlaAspSerArgProProAspGluIleYr 17
 |||||
 218 AATGGCGACAGATTATACCGTCTGACTGACCCAGATGAATAA 267
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPhaSpArg 34
 |||||
 268 ACGTTCGGAGGCTTATGCCAGAGGCCATATGAGTACTTGATAGAG 317
 34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||
 318 GAACCTAATGAATATTAATCTTATGATCACCAGAGACACACAAACC 367
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||
 368 GCGTTTGTCAGATGATGACGATATGTTCCACTTACCTTAGTTGAG 417
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 |||||
 418 AAGTGCTCAGCTTAGCAGACAGCTATATATATCAGATATCTCCACTTACT 467
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||
 468 ATATATATATGTTATAGCAGACGACCAATATGTTAATGTTAATGATGA 517
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
 |||||
 518 TTAGGCGTATACAGCCCTCACCATATGAACAGAGAGTTCTCGCTTAG 567
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAspPheGly 134
 |||||
 568 TGAATAACCATATTCAGATATATGATGATCGTGTAAATTTGCTG 617

134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||
 618 TGATGATGACAGATTACATCTGTAACAGGAAATATAGACCGGATATAC 667
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 |||||
 668 AGAAATCTGAAATATAGCTCCGCGAGAGATGGTTACAAATATAGCAGGTT 717
 167 eProProAspHisGlnAlaThrArgGluGluProProIleHisAlaP 184
 |||||
 718 CCACCGGATCACCAAGCTTGAGAGAAAGACCCTGGAATTCATTCGAC 767
 184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||
 768 CCAAGGCTTGGAATATCATCAGACATCAGAGGTATCTGTAT 817
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 |||||
 818 GAGAGAGACCCAGAACTGACCAATATATCTCAGAGGAATATCAATCAA 867
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 |||||
 868 AGTTAAGAGCGAGATATTTTCAGACTATCAGTACAGAGGTGACATATATA 917
 234 sNArgIleArgAspGluLeu 240
 |||||
 918 ACAGAAATTCGGATGATTA 937

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ23864

seq_documentation_block:

ID AAQ23864 standard; DNA; 1148 BP.

AC AAQ23864;

DT 08-OCT-1992 (first entry)

DE Thermolabile toxin (LTn) gene.

KW LTn; ST1a; ST1b; thermolabile; heat-stable; *E. coli*; T4DNA ligase;
 KW probe; detection; ss.

OS *Escherichia coli*.

PN JF04079898-A.

PD 13-MAR-1992.

XX 23-JUL-1990; 90JP-0194208.

XX 23-JUL-1990; 90JP-0194208.

XX (KITA) KITASATO RES INST.

DR WPI: 1992-137930/17.

XX DNA and RNA probe - for simultaneous detection of *Vibrio cholerae*
 PT and entero-toxicogenic *E. coli*

PS Claim 1; Page 1; 15pp; Japanese.

XX The thermolabile toxin (LTn) gene sequence is derived from
 CC enterotoxigenic *E. coli*. An LTn-ST1a-ST1b probe, wherein all
 CC all or part of the sequences of the three toxin genes represented
 CC in AAQ23864-66 are ligated to each other by T4DNA ligase, may be used
 CC for the detection of *Vibrio cholerae* and enterotoxigenic *E. coli*
 CC simultaneously in a simple way.
 CC The method has the same sensitivity as the dangerous radio-isotope
 CC detection method.

XX Sequence 1148 BP; 399 A; 205 C; 229 G; 315 T; 0 other:

alignment_scores:

Quality: 1292.00 Length: 240
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.750

alignment_block:

US-09-528-682-1 x AAQ23864 ..

Align seg 1/1 to: AAQ23864 from: 1 to: 1148

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspLysIleTy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 AATGGGACAAATATATACCGTCTGACTCTAGACCCCAATCAATATAA 104
17 SarGserGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 ACGTCCGGAGCTTATGCCCCAGAGGCGCATATGAGTACTCGATAGAG 154
34 LyrHrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyHrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GAACCTAAATGAATATTAATCTTATGATCAGCGAGAGAACACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GCGTTTGTCAGATATGATGACGATATGTTTCCACTTCTTGTGTTGAG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 AAGTGCCTACATTAGCAGACAGCTCTATATATCAGGATATTCACATTACT 304
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATATATATGTTATAGACAGACACAAATATGTTTAAATGTTATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGlyGlnGlyValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTAGGGGTATACAGCCCTCACCATATGAACAGAGGTTCTCGCTTAGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGGAAATACCATATTCAGATATATGATGATGATGCGTTAAATTTGGTG 454
134 alIleAspGlyArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGAATGATGACGATTACATCGTAACAGGGAATATAGAGCCGTTATPAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGAAATCTGAAATATAGCTCCGCGACAGAGATGTTACAGATTACAGGTTT 554
167 eProProAspHisGlnAlaLTPArgGlnGluProTyrPheHisAlaLar 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCACACCGGATACCAAGCTTGAGAGAAACCTGGATTCATCATGACAC 604
184 rGlnGlyCysGlyAsnSerSerArgThrIleHrGlyAspThrCysAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CACAAGGTTGTGAAATTCATCAAGAACAATTCAGGTGTTACTTGTAAAT 654
201 GlnGlyThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSerIly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAGGAGACCCAGAAATCTGAGCACAAATATATCTCAGCAAAATATCAACAA 704
217 sValLysArgGlnIlePheSerAspTyrGlnSerGlyValAspIleTyrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 AGTTAAGAGCAGATATTTTCAGACTATCATGTCAGAGGTTGACATATATA 754
234 snArgTLeaArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ACAGAAATTCGGAATGAATTA 774
```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS1106

seq_documentation_block:

ID AAAS1106 standard: cDNA; 782 BP.

XX AC AAAS1106;

XX 26-SEP-2000 (first entry)

DE Plant-optimized E. coli heat labile toxin A subunit coding sequence.

XX Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;

KW adjuvant; anti-bacterial; ss.

XX Escherichia coli.

OS Synthetic.

XX Key Location/Qualifiers

FH CDS 3..782

FT /*tag= a

FT sig_peptide 3..56

FT mat_peptide 57..779

FT /*tag= b

FT /*tag= c

PN WO200037609-A2.

XX 29-JUN-2000.

PD 22-DEC-1999; 99WO-US30747.

XX 22-DEC-1998; 98US-0113507.

PR (BOYC-) BOYCE THOMPSON INST PLANT RES.

XX (MASO/) MASON H S.

PA (ARNT/) ARNTZEN C J.

XX Mason HS, Arntzen CJ;

PI WPI: 2000-442653/38.

XX P-PSDB: AAY96646.

DR New polynucleotides encoding LT-A or CT-A polypeptides for the

XX transformation of plant cells; useful in immunogenic compositions to

PT elicit immune responses in animals

XX Example 1; Fig 1; 103pp; English.

PS This plant-codon optimized cDNA encodes a synthetic Escherichia coli

XX heat-labile toxin (LT) A subunit (LT-A). The sequence contains

CC plant-preferred codons and eliminates sequence motifs associated with

CC spurious mRNA processing. A single codon insertion (GTC encoding valine)

CC was made to accommodate the creation of a NcoI restriction site around the

CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A

CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit

XX (CT-A) polypeptide, which have reduced enzyme activity as compared to the

CC wild-type LT-A or CT-A polypeptide and where at least one of the codons

CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B

CC subunit (CT-B). The polynucleotides are useful for the transformation of

CC plant cells for the production of transgenic plants to produce edible

CC vaccines, especially oral vaccines in transgenic plants for the

CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The

CC mutant polypeptides are also useful as adjuvants.

XX

SQ Sequence 782 BP; 225 A; 186 C; 191 G; 180 T; 0 other;

alignment_scores:

Quality: 1287.00 Length: 240
Ratio: 5.362 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.333

alignment_block:

US-09-528-682-1 x AA51106 ..

Align seg 1/1 to: AA51106 from: 1 to: 782

```

1 ASnGLYAspArgLeuTyrArgAlaAspSerArgProAspGluIlely 17
  |||||||
60 AACGAGACAGACTCTATAGCGCTAGCGCCCAAGATGAGATCA 109
  |||||||
17 SARSGEGLYLYneumETProARGLYHISASnGLYTYrPheAspArg 34
  |||||||
110 GAGGTCTGAGAGTCTCATGCCAAGGACACAAATGACTTCTATAGG 159
  |||||||
34 LYthRGlnMetAsnIleAsnLeuTyrAspHisAlaArgGLYthRGlnThr 50
  |||||||
160 GAACCCAAATGAAATCATACCTCTATGACCATGCTAGGGGAAACCAACT 209
  |||||||
51 GlyPheValArgTyrAspAspGLYTYrValSerThrSerLeuSerLeuAr 67
  |||||||
210 GCGTTTGTGAGTGTGATGATGATGTCTCCACTCCCTTACTCTGAG 259
  |||||||
67 GSERAlAHISLeuAlaGLYInSerIleLeuSerGLYTYrSerThrTYrT 84
  |||||||
260 GTCTGCTCACTTGGCAGACATCCCTCTCTAGGATCTCCACTCTACT 309
  |||||||
84 YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||||||
310 ACATCTATGTGATGTCTACAGACACAAACATGTTCATGTGAATGATGTG 359
  |||||||
101 LeuGLYValTYrSerProHISProTYrGLInGLInValSerAlaLeuGL 117
  |||||||
360 TTGGAGATGTATAGCTCCTCACCATATGACCAAGAGTGTCTGTGGG 409
  |||||||
117 YGLYIleProTYrISerGLInIleTYrGLYTYrTYrArgValAsnPhelGLY 134
  |||||||
410 TGAATCCCATCTACTCCAAATCTATGTGATGTATAGGCTGAACTTGTGTG 459
  |||||||
134 alIleAspGluArgLeuHisArgAsnArgGLYTYrArgAspArgTYrTYr 150
  |||||||
460 TGATGTATGAGAGGCTCCATAGAGATAGGAGATATAGGAGCAGGTACTAT 509
  |||||||
151 ArgAsnLeuAsnIleAlaProAlaGLYAspGLYTYrArgLeuAlaGLYPh 167
  |||||||
510 AGGAACTCTCAACATAGCTCCAGCAGAGATGTATAGTGTGGCAGGT 559
  |||||||
167 eProProAspHISGLInAlaTrPARGLInGLInProTYrIleHISHisAlaP 184
  |||||||
560 CCCACAGACACACCAAGCTGAGAGGAGAGCCCTGATCCACCATGTCAC 609
  |||||||
184 roGLInGLYCYsGLYAsnSerSerArgThrIlethRGlnAspThrCysAsn 200
  |||||||
610 CACAAGGTTGTGAGACTCCTCAAGGACATCACAGGTGACACTTGCAAT 659
  |||||||
201 GLInGLYthRGlnAsnLeuSerThrIleTYrLeuArgGLInGLYInSerLY 217
  |||||||
660 GAGAGAGCCCAAAACCTTAGCACCATTTACCTTAGGAGATACCAATCAA 709
  |||||||
217 sValLYsArgGLInIlePheSerAspTYrGLInSerGLInValAspIleTYrA 234
  |||||||
710 GGTGAAGAGCAAAATCTTCTCAGACTACCAATCAGAGGTGAGACATCTACA 759
  |||||||
234 snArgIleArgAspGluLeu 240
  |||||||
760 ATAGATTAGGAATGAACCT 779

```

seq_name: /sids5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA51147

seq_documentation_block:

ID AA51147 standard; cDNA; 782 BP.

AC AA51147;

XX 26-SEP-2000 (first entry)

XX

DE Plant-optimized E. coli LT-A K63 mutant coding sequence.

KW Heat-labile toxin: LT-A; LT-B; mutant; transgenic plant; vaccine; oral;

KW adjuvant; anti-bacterial; S63K, ss.

OS Escherichia coli.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS 3..782

FT sig_peptide /tag= a

FT mat_peptide /tag= b

FT /tag= c

WO200037609-A2.

PD 29-JUN-2000.

PF 22-DEC-1999; 99WO-US30747.

PR 22-DEC-1998; 98US-0113507.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES.

PA (MASO/) MASON H S.

PA (ARNT/) ARNTZEN C J.

PI Mason HS, Arntzen CJ;

DR WPI: 2000-442653/38.

DR P-PSDB: AAY96647.

PT New polynucleotides encoding LT-A or CT-A polypeptides for the transformation of plant cells, useful in immunogenic compositions to elicit immune responses in animals

Example 2; Page -: 103pp; English.

This cDNA encodes a mutant S63K Escherichia coli heat-labile toxin (LT) A subunit (LT-A). The codon at nucleotide positions 246-248 was changed from the wild-type TCC to AAG which causes a change of serine to lysine in the mature protein at residue 63. The sequence contains plant-preferred codons and eliminates sequence motifs associated with spurious mRNA processing. A single codon insertion (GAG encoding valine) was made to accommodate the creation of a NcoI restriction site around the initiator methionine codon. Novel polynucleotides encode a mutant LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide, which have reduced enzyme activity as compared to the wild-type LT-A or CT-A polypeptide and where at least one of the codons is altered to a plant preferred codon. The polynucleotide further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The polynucleotides are useful for the transformation of plant cells for the production of transgenic plants to produce edible vaccines, especially oral vaccines in transgenic plants for the prophylactic or therapeutic treatment against E. coli or V. cholerae. The mutant polypeptides are also useful as adjuvants.

Note: This sequence does not appear in the specification. It was constructed from the wild type LT-A cDNA shown in AA51106 which is given in Figure 1 of the specification.

Sequence 782 BP; 227 A; 184 C; 192 G; 179 T; 0 other;

alignment_scores:

Quality: 1283.00

Ratio: 5.346

Percent Similarity: 100.000

Length: 240

Gaps: 0

Percent Identity: 97.917

alignment_block:

US-09-528-682-1 x AA51147 ..

Align seg 1/1 to: AAA51147 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleY 17
  |||
60 AACGGAGACAAAGCTCTATAGCGCTGACTAGGCCCCCAGATGATCA 109
  |||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  |||
110 GAGGTGTGAGGCTCTATGCGCAAGGGAGACAAATGAGTACTTGTGATG 159
  |||
34 1YThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
160 GAACCCAAATGAACATCAACCTCTATGACCATGCTAGGGGAACCAACT 209
  |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||
210 GGTCTGTGTGAGGTATGATGATGATGTGTCTCACCACCACTGACTTGA 259
  |||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||
260 GTCTGCTCACTTGGCAGACAAATCCATCCTCTCAGGATCTCCACTACT 309
  |||
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
310 ACATCTATGTGATGTCTACAGCACCAACATGTTCATGTGATGATGTG 359
  |||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
  |||
360 TTGGGAGGTATAGCCCTCACCCTATGTAGCAAGAGTGTCTCTTTGGG 409
  |||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
  |||
410 TGAATGCCATCTCCCAATCTATGATGATGTAGGCTGAACCTTGGTG 459
  |||
134 aIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
  |||
460 TGATTGATGAGAGGCTCCATAGGAATAGGAGATAGGGGACAGTACTAT 509
  |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  |||
510 AGGAACCTCAACATAGCTCCACGACGAGATGGTATAGGTTGGCAGGTT 559
  |||
167 eProProAspHisGlnAlaThrArgGluGluProTyrIleHisHisAla 184
  |||
560 CCCACCAACACCAACCTCGAGGAGAGAGCCCTGATCCACCAATGAC 609
  |||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||
610 CACAAGGTTGTGAGACTCCTCAAGACCATCAGGTGACTTGCAT 659
  |||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  |||
660 GAGGACACCAAAACCTTAGCACCATCTACCTTAGCAAGACCAATCCAA 709
  |||
217 sValIleAspArgGlnIlePheSerAspTyrGlnSerGluValAspIle 234
  |||
710 GGGAGAGAGCAATCTCTCAAGACTACCAATCAGAGGTGACATCTACA 759
  |||
234 snArgIleAspArgGluLeu 240
  |||
760 ATAGGATTAGGAATGAATC 779

```

seq_name: /SIDS5/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AAA51545

seq_documentation_block:

ID AAA51545 standard; cDNA: 782 BP.

AC AAA51545;

DT 26-SEP-2000 (first entry)

DE Plant-optimized E. coli LT-A R72 mutant coding sequence.

```

KW Heat-labile toxin: LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
KM adjuvant; anti-bacterial; A72R; SS.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT CDS 3..782
FT sig_peptide /tag= a
FT sig_peptide /product= heat-labile_toxin-A
FT mat_peptide /tag= b
FT /tag= c
PN WO200037609-A2.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US30747.
XX
XX 22-DEC-1998; 98US-0113507.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
XX (MASO/) MASON H S.
XX (ARNT/) ARNTZEN C J.
XX
XX Mason HS, Arntzen CJ;
XX
XX WPI: 2000-442653/38.
XX DR P-PSDB: AAY96649.
XX
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
XX transformation of plant cells, useful in immunogenic compositions to
XX elicit immune responses in animals
XX
XX Example 4; Page -: 103pp; English.
XX
XX This cDNA encodes a mutant A72R Escherichia coli heat-labile toxin (LT)
XX A subunit (LT-A). The nucleotides at position 273-274 were changed from
XX CC the wild-type GC to AG which causes a change of alanine to arginine in
XX CC the mature protein at residue 72. The sequence contains plant-preferred
XX CC codons and eliminates sequence motifs associated with spurious mRNA
XX CC processing. A single codon insertion (GTC encoding valine) was made to
XX CC accommodate the creation of a NcoI restriction site around the
XX CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
XX CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
XX CC (CT-A) polypeptide, which have reduced enzyme activity as compared to
XX CC the wild-type LT-A or CT-A polypeptide and where at least one of the
XX CC codons is altered to a plant preferred codon. The polynucleotide further
XX CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
XX CC subunit (CT-B). The polynucleotides are useful for the transformation of
XX CC plant cells for the production of transgenic plants to produce edible
XX CC vaccines, especially oral vaccines in transgenic plants for the
XX CC prophylactic or therapeutic treatment against E. coli or V. cholerae.
XX CC The mutant polypeptides are also useful as adjuvants.
XX CC Note: This sequence does not appear in the specification. It was
XX CC constructed from the wild type LT-A cDNA shown in AAA51106 which is
XX CC given in Figure 1 of the specification.
XX
XX Sequence 782 BP; 226 A; 185 C; 191 G; 180 T; 0 other;
XX

```

alignment_scores:

Quality: 1282.00 Length: 240
Ratio: 5.364 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 97.917

alignment_block:

US-09-528-682-1 x AAA51545 ..

Align seg 1/1 to: AAA51545 from: 1 to: 782

XX	Escherichia coli.
OS	Synthetic.
XH	Key
FH	Location/Qualifiers
FT	CDS
FT	3..782
FT	/tag- a
FT	/product= heat-labile-toxin-A
FT	sig_peptide
FT	3..59
FT	/tag- b
FT	mat_peptide
FT	60..779
XX	/tag- c
PN	W0200037609-A2.
XX	
PD	29-JUN-2000.
XP	
XP	22-DEC-1999; 99WO-US30747.
XP	
PR	22-DEC-1998; 98US-0113507.
PA	(BOYCE-) BOYCE THOMPSON INST PLANT RES.
PA	(MASON/) MASON H S.
PI	(ARNTZ/) ARNTZEN C J.
PI	Mason HS, Arntzen CJ;
XX	
DR	WPI: 2000-442653/38.
DR	P-PDB; AAY96648.
PT	New polynucleotides encoding LT-A or CT-A polypeptides for the
PT	transformation of plant cells, useful in immunogenic compositions to
PT	elicit immune responses in animals
PS	Example 3; Page -; 103pp; English.
XX	
CC	This CDNA encodes a mutant R192G Escherichia coli heat-labile toxin (LT)
CC	A subunit (LT-A). The nucleotide at position 631 was changed from the
CC	wild-type A to G which causes a change of arginine to glycine in the
CC	mature protein at residue 192. The sequence contains plant-preferred
CC	coding and eliminates sequence motifs associated with spurious mRNA
CC	processing. A single codon insertion (GTG encoding valine) was made to
CC	accommodate the creation of a NcoI restriction site around the
CC	initiator methionine codon. Novel polynucleotides encode a mutant LT-A
CC	polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
CC	(CT-A) polypeptide, which have reduced enzyme activity as compared to
CC	the wild-type LT-A or CT-A polypeptide and where at least one of the
CC	codons is altered to a plant preferred codon. The polynucleotide further
CC	comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC	subunit (CT-B). The polynucleotides are useful for the transformation of
CC	plant cells for the production of transgenic plants to produce edible
CC	vaccines, especially oral vaccines in transgenic plants for the
CC	prophylactic or therapeutic treatment against E. coli or V. cholerae.
CC	The mutant polypeptides are also useful as adjuvants.
CC	Note: This sequence does not appear in the specification. It was
CC	constructed from the wild type LT-A CDNA shown in AAA51106 which is
CC	given in Figure 1 of the specification.
XX	
Sequence 782 BP; 224 A; 186 C; 192 G; 180 T; 0 other;	

seq_documentation_block:

AC AAA51544;

DT 26-SEP-2000 (first entry)

Plant-optimized E. coli LT-A G192 mutant coding sequence

KM Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral
KM adjuvant; anti-bacterial; R192G; ss.

alignment_scores:

Quality:	1280.00	Length:	240
Ratio:	5.356	Gaps:	0
Percent Similarity:	99.583	Percent Identity:	97.917

```
alignment_block:
```

Align seg 1/1 to: AAA51544 from: 1 to: 782

1 ASNGlyaspArgLeuTyrArgAlaaspSerArgProProaspGluIleIleY 17


```

60 AACGAGACAGCTCTATAGGCTGACTCTAGGCCCCAGATGATCA 109
17 sarserglyglyleumetproarglyhisnglutyrpheasparg 34
110 GAGGTCTGGAGGTCTCATGCCAAGGGGACAAATGAGTCTTGATGAGG 159
34 lythrlmetasnileasnleuetyrasphialaarglythrlnthr 50
160 GAACCCAAATGAACATCAACCTCTATAGACCATCTGAGGAAACCAACT 209
51 glyphevalargtyraspspelytyrvalserthrserleuaur 67
210 GGCTTGATGAGGTATGATGATGATGATGATGATGATGATGATGATG 259
67 gseralahlisleualaglycinsertleuenserglytyrserthtyt 84
260 GTCTGCTCACTTGGCAGGACAAATCCATCCTCAGATACTCCACTACT 309
84 yrlletyrvalilealathralaproasnmetpheasnvalasnspval 100
310 ACATCTATGTGATGTCTACAGACCAACATGTTCAATGTGATGATG 359
101 leuglyvaltyrserprohisprotyrgluingluvalseralaleucl 117
360 TTGGGAGGTATAGCCCTCACCATATGAGCAAGAGTGTCTGCTTGGG 409
117 yglylieprotyrserglniletyrcltyrtyrargvalasnphgly 134
410 TGGAAATCCATCTACCCAAATCTATGATGATGATGATGATGATGATG 459
134 alileaspluarlgleunhisarasnarglytyrargspargtyrtyr 150
460 TCAATCATATAGAGGCTCCATAGGAATAGGAGTATAGGACAGTACTAT 509
151 argasnleuasnllealaproalaagluaspglytyrargleualaagly 167
510 AGGAACCTCAACATAGCTCCAGACAGAGATGATGATGATGATGATG 559
167 eproproasphlsglnaltrparglugluprotprllehlshlsalpr 184
560 CCCACACAGACCAACCAAGCTGGAGGAGAGGCCCTGGATCCACCAATG 609
184 roglnglycysglasnserserargthrilethrlyasprthrcysasn 200
610 CACAAGGTGTGAGACTCTCAGGACCATCAACAGCTACACTTCAAT 659
201 gluglutthrghlnasnleuserthriletyrleuarglytyrclnsery 217
660 GAGGAGACCCAAACCTTAGCACATCTAGCTTAGGAGTACCAATCA 709
217 slyllysatrglnilepheseraptyrclnsertgluvalaspilertyr 234
710 GGTGAAGAGGCAAAATCTCTCAGACTACCAATCAGAGTGCATCTACA 759
234 snargtleargaspgluleu 240
760 ATAGGATTAGGAATGACTC 779
seq_name: /SIDS5/gcdata/geneseq/geneseq-emb1/NA2001A.DAT: AAS01506
seq_documentation_block:
ID AAS01506 standard; DNA; 1508 BP.
AC
XX
AS01506;
XX
29-AUG-2001 (first entry)
DE
XX
E. coli heat-labile enterotoxin (LT) mutant ltrdel110/112 DNA.
XX
Heat-labile enterotoxin; LT; LTrS63Y; LTrdel110/112; mutant;
KW detoxified and immunologically active protein; ADP-ribosylation; Gs;
endotoxin; diarrhoea; ds.
XX

```

```

OS Escherichia coli strain K88ac.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 164..1305
FT
FT     /tag="a
FT     /product="Ltrdel110/112 mutant protein"
FT     /trans_except="(pos:932..933,aa:Met)
FT     /note="This codon has an apparent 1 nucleotide
FT     deletion which alters the reading frame"
XX
XX MO200119998-A1.
XX
XX 22-MAR-2001.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX
XX WPI: 2001-281524/29.
XX
XX P-PSDB: AAU00507.
XX
XX New detoxified mutants of Escherichia coli heat-labile enterotoxin
XX useful as vaccine for preventing and treating diarrhoea, and as adjuvant
XX for antibody production.
XX
XX
XX Claim 7; Page 45-46; 48pp; English.
XX
XX
XX The present sequence encodes for Escherichia coli heat-labile
XX enterotoxin (LT) mutant Ltrdel110/112. LTrS63Y (AAU00506) and Ltrdel110/112
XX are two novel detoxified and immunologically active proteins (LT
XX mutants) derived by site-directed mutagenesis of the A1 subunit of wild
XX type LT. The substitution of Ser to Tyr at position 63 in LTrS63Y blocks
XX NAD-binding. Deletion of Glu residues at positions 110 and 112 in
XX Ltrdel110/112 eliminate the enzymatic activity of LT. The A1 subunit of
XX wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
XX regulates cAMP levels. The resulting increase in cAMP is the cause of
XX diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
XX of mutant heat-labile enterotoxins LTrS63Y and Ltrdel110/112 were tested.
XX Groups of mice were immunised with LTrS63Y or Ltrdel110/112. The control
XX groups received phosphate buffered saline (PBS) alone. The serum and
XX faecal antibody titres to LT were determined. The results showed that
XX mice immunised with LTrS63Y or Ltrdel110/112 contained high and
XX comparable level of anti-LT antibodies in sera and faecal extracts
XX compared with those immunised with wild-type LT. The LT mutants are
XX useful as a vaccine for preventing and treating diarrhoea and as an
XX adjuvant for antibody production.
XX
XX Sequence 1508 BP; 506 A; 277 C; 295 G; 430 T; 0 other:
SQ

```

```

alignment_scores:
Quality: 1280.00 Length: 240
Ratio: 5.378 Gaps: 1
Percent Similarity: 99.167 Percent Identity: 98.750

```

```

alignment_block:
US-09-528-682-1 x AAS01506 ..

```

```

Align seg 1/1 to: AAS01506 from: 1 to: 1508

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspLysLeu 17
|||||
218 AATGGCGACAGATTATACCGTGTGACTCTAGACCCCGAGATAAATAA 267
17 sarserglyglyleumetproarglyhisnglutyrpheasparg 34
|||||
268 ACCTCCGAGAGTCTATAGCCAGAGGCATATAGTACTTGATAGAG 317

```



```

34  lYthrGlmetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
318  GAACTCAATGATATATATCTTATGATCAGCGAGAGAACACAAACC 367
51  GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
368  GCGTTTGTCAGATATGATGACGATATGTTCCACTTCTTAGTTGAG 417
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
418  AAGTCTCACTAGCAGACAGACTATATATATCAGATATTCCTACTTACT 467
84  yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
468  ATATATATGTTATGACGACAGACCAATATGTTTAAATGTTATGATGTA 517
101  LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG1 117
518  TTAGGCGATATACAGCCCTCACCATAT.....CAGGTTTCGCGTTAGG 561
117  yGlyIleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPhgGlyV 134
562  TCGATATACATATCTTCAGATATATGAGATGATCTGTTATTTTGTG 611
134  aIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
612  TGATTGATGATACGATTTACATCGTAACAGCGAATATAGAACCGGTAATAC 661
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyph 167
662  AGAAATCTGAATATAGCTCCGACAGAGATGTTACAGATTAGCAGGTTT 711
167  eProProAspHisGlnAlaTyrPArgGluGluProTyrIleHisAlaP 184
712  CCCACCGGATTCACCAAGCTTGAGAGAAACCCGATTCATCATGCAC 761
184  roGlnGlyCysGlnSerSerArgThrIleThrGlyAspThrCysAsn 200
762  CACAAGGTTGTGAAATTCATCAGAACATCAGAGGTGATCTGTATAT 811
201  GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerY 217
812  GAGGAGACCCAGAAATCTGAGCAAAATATCTCAGGGAAATATCAATCAA 861
217  sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
862  AGTTAAGAGCGAGATATTTTCAGACATCATCAGACAGGTTTGACATATA 911
234  snArgIleArgAspGluLeu 240
912  ACAGAAATTCGGGATGAAATTA 931
seq_name: /SIDS5/gcgcdata/geneseq/geneseq-emb1/NA1993.DAT.AA042768
seq_documentation_block:
ID AA042768 standard; DNA, 711 BP.
XX
AC AA042768;
XX
DT 08-DEC-1993 (first entry)
XX
DE E.coli heat labile toxin subunit A coding sequence.
XX
KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.
XX
OS Escherichia coli.
XX
FH Key location/Qualifiers
FT mat_peptide 1..708
FT /*tag= a
FT /note= "LT-A"

```

```

XX
PN WO9313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91IT-0M13513.
XX
PA (BIOC-) BIOCINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Pizza M, Rappuoli R;
XX
DR WPI: 1993-227320/28.
XX
DR P-PSDB; AAR38728.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
PS Disclosure; Fig 2; 60pp; English.
XX
CC This sequence encodes the A subunit of the heat labile toxin (LT-A)
CC of a strain of E.coli known to affect humans. The sequence was
CC published by Yamamoto et al, J.Biol. Chem., 259, 5037-5044.
CC Mutations at selected positions within this sequence have been found
CC to reduce toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli.
XX
SQ Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other:

```

```

alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333

```

```
alignment_block:
US-09-528-682-1 x AA042768 ..

```

```
Align seg 1/1 to: AA042768 from: 1 to: 711

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleY 17
|||||
1 AATGGCAGACAGATTAATACCGTCTGACTGATAGACCCAGATGAATAAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlnIleTyrPheAspArg 34
|||||
51 ACGTTTCGCGAGCTTATGCCAGAGGT...AATGATGACTTCGATATAG 97
34 lYthrGlmetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
98 GAACTCAAATGATATATATCTTATGATCAGCGACGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
148 GCGTTTGTCAGATATGATGACGATATGTTCCACTTCTTAGTTGAG 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
198 AAGTCTCACTAGCAGACAGACTATATATATATCAGATATTCCTACTTA 247
84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
248 TATATATCTGTTATGCA.....AATATGTTTAAATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG1 117
|||||
289 ATTAGCGATATACAGCCCTCACCATATGACAGAGAGGTTTCGCGTTAG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPhgGlyV 134
|||||

```



```

339 TGAATACCAATATCTCAGATATATGATGGTATCGTGAATATTTGGTG 388
134 allleasrgluatrgleuHisargAsnargluYrArgaspargYrTyr 150
389 TGAATGATGAACGATTCATCTGTAACAGGAAATATAGACCGGTATATAC 438
151 ArgAsnleuAsnIleAlaProAlaGluAspGlyTyrArgleuAlaGlyPh 167
439 AGAATATGATATAGCTCCGGCAGAGAGATGGTTACAGATTAGCAGGTTT 488
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisIAsIaP 184
489 CCACCGCGATACCAACCTTGAGAGAGAAACCTGATTCATCATATCAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAGAGTTGTGGAGATTCAATCAAGAAACATCACAGCTGATCTTGAAT 588
201 GluGluThrGlnAsnleuSerThrIleTyrleuArgGluYrGlnSerly 217
589 GAGAGACCCAGAACTTGACACAAATATCTCAGGAAATATCAATCAA 638
217 sVallyArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
639 AGTTAAGACGACAGATATTTTCAGACATCATGTCAGAGGTTGACATATA 688
234 sNaArgIleArgAspGluLeu 240
689 ACAGAAATCGGAGTAATTA 708

```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV81595

seq_documentation_block:

ID AAV81595 standard; DNA; 711 BP.

AAV81595;

11-MAY-1999 (first entry)

E. coli heat labile toxin gene.

A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;
 parental adjuvant; antigen; antitumor; immunisation; humoral response;
 cell-mediated immune response; virus; bacterium; parasite; fungus;
 tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;
 systemic lupus erythematosus; Alzheimer's disease; diagnosis; ss.

Escherichia coli.

MO9842375-A1.

01-OCT-1998.

19-MAR-1998; 98WO-US05454.

18-MAR-1998; 98US-0044696.

21-MAR-1997; 97US-0041227.

(CHIR) CHIRON CORP.

Barchfeld G, Del Giudice G, Rappuoli R;

WI: 1999-070064/06.

P-PSDB; AAM67772.

Detoxified mutants of bacterial ADP-ribosylating toxins as
 parental adjuvants - useful to enhance humoral and cell-mediated
 immune responses in vertebrates when administered with selected
 antigen e.g. in disease treatment

Disclosure: Fig 1A-B; 51pp; English.

This sequence corresponds to the coding region for the A subunit of the

CC E. coli heat labile toxin, an example of a bacterial ADP-ribosylating
 CC toxin. A mutant detoxified form of this protein is used in a parenteral
 CC adjuvant composition, which comprises the detoxified protein, at least
 CC one selected antigen and optionally a pharmaceutically acceptable
 CC (optionally topical) vehicle. The adjuvant composition can be
 CC administered parenterally in conjunction with at least one antigen in
 CC methods to immunise vertebrate subjects. The adjuvant has the ability
 CC to enhance the humoral and cell-mediated immune responses elicited by
 CC the antigen (e.g. by making the antigen more strongly immunogenic or
 CC necessitating fewer/lower antigen doses). It can be administered
 CC prior/subsequent to the antigen, and is preferably administered within
 CC a short space of time to the same site; it can also be administered in
 CC isolation from antigens as a boost following systemic or mucosal antigen
 CC administration. Most preferably, the adjuvant is co-administered with
 CC the antigen in the compositions and a pharmaceutically acceptable
 CC carrier. The antigen may be derived from viruses, bacteria, parasites
 CC and fungi or may be tumour antigens, self-antigens and allergens. The
 CC compositions are therefore useful in the treatment and prevention of
 CC e.g. viral diseases, allergic manifestations, diseases caused by
 CC pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases
 CC (e.g. Systemic Lupus Erythematosus), Alzheimer's disease and cancers.
 CC The adjuvant can also be used to prepare antibodies against selected
 CC antigen(s), useful e.g. for diagnostic purposes or for antigen
 CC purification.

Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other;

alignment_scores:

Quality: 1183.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:

US-09-528-682-1 x AAV81595 ..

Align seg 1/1 to: AAV81595 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
1 AATGGCAGACAGATATATACCGCTGACTCTAGACCCCAATGAAATAA 50
17 sArgSerGlyGlyleuSerProArgGlyHisAsnGlyTyrPheAspArg 34
51 ACGTTCCGAGGCTGTATGCCAGAGGT...AAGAACTACTTCGATGAG 97
34 lYThrGlnMetAsnIleAsnleuTyrAspHisAlaAspGlyThrGlnThr 50
98 GAACTCAATGAATATATATCTTATATGATCAGCGAGAGAGAACCAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerleuSerleuAr 67
148 GCGTTTGTCAAGATATGATGACGAGATATGTTCACCTCTCTAGTTGAG 197
67 gSerAlaHisIleuAlaGlyGlnSerIleleuSerGlyTyrSerThrTyrT 84
198 AAGTGTCACTTAGCAGACAGATATATATATATCAAGATATCACTTACA 247
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATAGCA.....AATATGTTTAAATGATGATGA 288
101 leuGlyValTyrSerProHisProTyrGluGlnIleValSerAlaLeuG 117
289 ATTAGCGTATACAGCCCTCAACATATGAAACAGAGGCTTCTCGCTTAG 338
117 uGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAspPheGly 134
339 TGAATATCAATATCTCAGATATATGATGATCGATTAAATTTGGTG 388
134 allleasrgluatrgleuHisargAsnargluYrArgaspargYrTyr 150
389 TGAATGATGAACGATTCATCTGTAACAGGAAATATAGACCGGTATATAC 438

```



```

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
167 ePProAspHisGlnAlaIleAlaTyrPargLugluPProTrrPleHisAlaPhe 184
489 AGAAATCTGAATATAGCTCCGCGAGAGAGATGTTACAGATTACAGAGTTT 488
167 ePProAspHisGlnAlaIleAlaTyrPargLugluPProTrrPleHisAlaPhe 184
489 CCCACCGGATCCACCAAGCTTGAGAGAGAAACCTGGATTCCATCATGTCAC 538
184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTTGAGATTTCATCAAGAACATTCACAGGTGATTGTTAAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLys 217
589 GAGGAGACCCAGATCTGACACACAATATCTCAGGSAATATCAATCAAA 638
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
639 AGTTAAAGCGCAGATATTTTCAGACTATCAGTCAGAGGTTGACATATATA 688
234 sNArgIleArgAspGluLeu 240
689 ACAGAAATTCGGAGATGATTA 708
seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051317

```

```

seq_documentation_block:
ID AA051317 standard; DNA; 711 bp.

```

```

XX AA051317;

```

```

DT 08-DEC-1993 (first entry)

```

```

DE Encodes Lys-63 E.coli heat labile toxin subunit A.

```

```

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.

```

```

XX Escherichia coli.

```

```

XX Key Location/Qualifiers
FT misc_difference 184..186

```

```

FT /tag= a
FT /note= "wild-type TCT(Ser) mutated to AAG(Lys)"

```

```

XX WO9313202-A.

```

```

XX 08-JUL-1993.

```

```

XX 30-DEC-1992; 92WO-EP03016.

```

```

XX 31-DEC-1991; 91IT-OMI3513.

```

```

XX (BIOC-) BIOCINE SCLAVO SPA.

```

```

XX Domenighini M, Hol W, Pizsa M, Rappuoli R;

```

```

XX MPI: 1993-227320/28.

```

```

XX P-PSDB; AAR44016.

```

```

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

```

```

XX Claim 3; Fig 2 and Page 46; 60pp; English.

```

```

XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-Q51326). The invention relates to

```

```

CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051317 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 bp; 236 A; 122 C; 157 G; 196 T; 0 other;

```

```

alignment_scores:
Quality: 1179.00 Length: 240
Ratio: 5.171 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 92.917

```

```

alignment_block:
US-09-528-682-1 x AA051317

```

```

Align seg 1/1 to: AA051317 from: 1 to: 711

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLys 17
1 AATGGCGACAGATTATACCGTGCTGACTCTAGACCCCGAGATGAATATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 ACCTTCCGAGCTTATGCCAGAGGT...AATGAGTACTTCGATAGAG 97
34 IYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
98 GAACCTAAATGAATATTAATCTTTATGATCAGCCGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
148 GCGCTTGTCAGATATGATGATGCGGATATGTTCCACTAAGCTTAGTTGAG 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
198 AAGTGTCTCCTTATGCGAGACAGATATATATATATATATATATATATAT 247
84 yrlLetyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATAGCA.....AATATGTTTATGTTATATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGlnGluGluValSerAlaLeuG 117
289 AATAGCGTATACAGCCCTCAGATATGAACAGAGAGTTTCTCGTTAG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
339 TGGAAATACCATATTCAGATATATATGATGATGATGATGATGATGATG 388
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
389 TGAATTATGACCATTTACATCGTAACAGGAAATATAGAACCGGTATTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
439 AGAAATCTGAATATAGCTCCGCGAGAGAGATGTTACAGATTACAGAGTTT 488
167 ePProAspHisGlnAlaIleAlaTyrPargLugluPProTrrPleHisAlaPhe 184
489 CCCACCGGATCCACCAAGCTTGAGAGAGAAACCTGGATTCCATCATGTCAC 538
184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTTGAGATTTCATCAAGAACATTCACAGGTGATTGTTAAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLys 217
589 GAGGAGACCCAGATCTGACACACAATATCTCAGGSAATATCAATCAAA 638
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234

```


639 AGTTAAGAGGACAGATATTTTCAGACTATCAGTCAGAGGTTGACATATATA 688

234 snArg11leArgaspGluLeu 240

689 ACAGAAATTCGGGATGAAATTA 708

seq_name: /SID5/gcgdata/geneseq/geneseqn-embL/NA1993.DAT:AAQ51325

seq_documentation_block:

ID AAQ51325 standard; DNA: 711 BP.

AAQ51325;

08-DEC-1993 (first entry)

Encodes Glu-114 E.coli heat labile toxin subunit A.

enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

protomer A; site-directed mutagenesis; reduced toxicity;

ADP-ribosyltransferase activity; ss.

Escherichia coli.

Key Location/Qualifiers

misc_difference 328..330

08-JUL-1993.

30-DEC-1992: 92WO-EP03016.

31-DEC-1991: 91IT-OMI3513.

(BIOC-) BIOCLINE SCLAVO SPA.

Domenighini M, Hol W, Piazza M, Rappuoli R;

WP1; 1993-227320/28.

P-PSDB; AAR44024.

Immunogenic detoxified mutant cholera toxin and heat labile toxin

- useful as vaccines against infection by Vibrio cholerae and

enterotoxin producing Escherichia coli

Claim 3: Fig 2 and Page 46; 60pp; English.

The wild-type sequence coding for the A subunit of the heat labile toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AAQ51314-051326). The invention relates to immunogenic, detoxified LT-A proteins and their use in vaccines to protect against enterotoxigenic E.coli. Sequence AAQ51325 is a combination of the wild-type coding sequence and the mutagenic primer sequence used to introduce the preferred mutation. (Amino acid numbering is based on the cholera toxin A subunit sequence).

Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

Alignment_scores:

Quality: 1179.00

Ratio: 5.171

Percent Similarity: 95.000

alignment_block:

US-09-528-682-1 x AAQ51325 ..

Length: 240

Gaps: 2

Percent Identity: 92.917

Align seg 1/1 to: AAQ51325 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17

1 AATGGGACGATATATACCGTCTGACTAGACCCACATGAAATATA 50

17 sarGserGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34

51 ACCTTCCGGAGTCCTATGCCAGAGGT...AATGACTACTTGATAGAG 97

34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50

98 GAATCAAAATGAATATTAATCTTATGATCAGCGAGAGGACCAAAACC 147

51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67

148 GGCTTGTGCAGATATGATGAGGAGTATGTTCCACTCTCTGTAGTTGAG 197

67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrTr 84

198 AATGGCTCACTTAGCAGACAGATATATATATCAGGATATCACTTACTTA 247

84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100

248 TATATATCGTTATATAGCA.....AATATGTTATATGTTATATGATGTA 288

101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117

289 AATAGGCTATACAGCCCTCACCCATATGACAGAGAGGTTCAGCGTTAGG 338

117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134

339 TGGATATACCATATTCAGATATATGATGATCGTAAATTTTGCTG 388

134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150

389 TGATTGATGAACGATATACATCGTAACAGCGAATATATAGACCCGTAATTAC 438

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167

439 AGAAATCTGAATATATGCTCCGCGACAGAGATGTTACGATTACAGGTTT 488

167 eProProAspHisGlnAlaTyrArgGluGluProTyrIleHisIleAla 184

489 CCCACCGGATCCACCAAGCTTGGAGAGAGAACCTGGATTCATCATGAC 538

184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200

539 CACAAGGTTGTGAGATTCATCAAGAACATACAGAGTGATCTGTAAT 588

201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217

589 GAGGAGACCCAGAAATCTGACACAAATATATCTCAGGAAATATCAATCAA 638

217 sValIlySarGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234

639 AGTTAAGAGGACAGATATTTTCAGACATATCAGTCAGAGGTTGACATATATA 688

234 snArg11leArgaspGluLeu 240

689 ACAGAAATTCGGGATGAAATTA 708

seq_name: /SID5/gcgdata/geneseq/geneseqn-embL/NA1993.DAT:AAQ51326

seq_documentation_block:

ID AAQ51326 standard; DNA: 711 BP.

AAQ51326;

08-DEC-1993 (first entry)

Encodes Lys-114 E.coli heat labile toxin subunit A.

enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

protomer A; site-directed mutagenesis; reduced toxicity;


```

KM ADP-ribosyltransferase activity; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_difference 328..330
FT /*tag= a
FT /note= "wild_type TCT(Ser) mutated to AAA(Lys)"
XX
PN MO9313202-A.
XX
XX 08-JUL-1993.
XX
XX 30-DEC-1992; 92MO-EP03016.
XX
XX 31-DEC-1991; 91IR-0M13513.
XX
XX (BIOC-) BIOCIINE SCLAVO SPA.
PI Domenighini M, Hol W, Piazza M, Rappuoli R;
XX
XX WPI: 1993-227320/28.
XX P-PSDB: AAR44025.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX - useful as vaccines against infection by Vibrio cholerae and
XX enterotoxin producing Escherichia coli
XX
XX Claim 3; Fig 2 and Page 46; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the heat labile
XX toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
XX et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
XX site-directed mutagenesis. Certain mutations were found to reduce
XX toxicity (see AA051314-Q051326). The invention relates to
XX immunogenic, detoxified LT-A proteins and their use in vaccines to
XX protect against enterotoxigenic E.coli. Sequence AA051326 is a
XX combination of the wild-type coding sequence and the mutagenic
XX primer sequence used to introduce the preferred mutation. (Amino
XX acid numbering is based on the cholera toxin A subunit sequence).
XX
XX Sequence 711 BP: 237 A; 122 C; 156 G; 196 T; 0 other;
XX
XX
XX alignment_scores:
XX Quality: 1179.00 Length: 240
XX Ratio: 5.171 Gaps: 2
XX Percent Similarity: 95.000 Percent Identity: 92.917
XX
XX alignment_block:
XX US-09-528-682-1 x AA051326 ..
XX
XX Align seg 1/1 to: AA051326 from: 1 to: 711
XX
XX 1 AAsnGIyAspArgLeuTYrARgAlaAspSerARgProProAspGluIleLY 17
XX |||||||.....
XX 1 AATGGGACAGATTTATACCGTGTCTACGCTAGACCCCGACAGATGAATATAA 50
XX |||||||.....
XX 17 sARgSerGIyGIyLeuMetProARgGIyHisAsnGIuTYrPhesARgG 34
XX |||||.....
XX 51 AAGTTCGCGAGCTCTTATGCCAGAGGT...AATGAGTACTTCCGATAGAG 97
XX |||||.....
XX 34 LYThGIhMeTAsnIleAsnLeuTYrASpHisAlARgGIyThGIhPthr 50
XX |||||||.....
XX 148 GACATCCAAATGAAATTTAATCTTTTATGATCACGCCAGAGGAAACACAAAC 147
XX |||||||.....
XX 51 GlyPheValARgTYrAspAspGIyTYrValSerThrSerLeuSerLeuAR 67
XX |||||||.....
XX 67 gSerAlaHisLeuAlaGIyGlnSerIleLeuSerGIyTYrSerThrTYT 84
XX |||||||.....
XX 198 AAGTGTCTACTTAGCAGGACGATATATTTATTCAGATATTCCTACTACTA 247

```

[illegible]

PI Domenighini M, Hol W, Pizze M, Rappuoli R;
 XX WPI: 1993-227320/28.
 DR P-PSDB: AAR38732.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PR - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli

PS Claim 3: Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051316 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP, 235 A; 123 C; 155 G; 198 T; 0 other;

alignment_scores:

Quality: 1178.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
 US-09-528-682-1 x AA051316 ..

Align seg 1/1 to: AA051316 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleTy 17
 1 AATGGGACAGATTATACCGCTGACTGACGCCCAATGATAATAA 50
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 51 ACCTTTCGGAGCTTATGCCAGAGGT...AATGAGTACTGCGATAGAG 97
 34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 98 GAACCTAATGATATTATATCTTTATGATCAGCGAGAGAACCAACACC 147
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
 148 GCGTTTACGATATGATGCGATATGTTCCACTTCCTTGTGTTGAG 197
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 198 AACTGCTCACTAGCAGACAGATATATATCAGGATATTCACCTTACTA 247
 84 YrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 248 TATATTCGTTATAGCA.....AATATGTTAATGTTAATAGATGA 288
 101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 289 ATTAGGCTATACAGCCCTCCATATGACAGAGAGTTCTCGCTTAGG 338
 117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 339 TGGAAATACCATATTCTCAGATATATGATGATGCGTAAATTTGGTG 388
 134 alIleAspGluArgLeuHisArgAsnArgIleTyrArgAspArgTyrTyr 150
 389 TGAATGAGCAATACATCGTACAGGGAATATATAGACCGGATATAC 438
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167

439 AGAATCTGAATATAGCTCCGACAGAGATGTTTACGATTACAGGTTT 488
 167 ePrProAspHisGlnAlaIleTyrArgGluGluProTyrIleHisIleAla 164
 489 CCCACCGGATCACCAGCTTGGAGAGAAACCTGGATTTCATCATGCAC 538
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 539 CACAGGTTTGAGATTCATCAAGAAACATCAGAGTGTACTTGTAAAT 588
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerTy 217
 589 GAGGAGACCCAGAAATCTGACACAAATATATCAGGAAATATCAATCAA 638
 217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 639 AGTTAAAGACAGATATTTTCACATCATCATCAGAGGTTGACATATATA 688
 234 snArgIleArgAspGluLeu 240
 689 ACAGAAATTCGGGATGAAATTA 708

seq_name: /SIDS5/gcgsdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051319

seq_documentation_block:

ID AA051319 standard; DNA; 711 BP.
 AC AA051319;
 DT 08-DEC-1993 (first entry)
 DE Encodes Tyr-97 E.coli heat labile toxin subunit A.
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
 KW protomer A; site-directed mutagenesis; reduced toxicity;
 KW ADP-ribosyltransferase activity; ss.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT misc_difference 277..279 /tag= a
 FT /note= "wild-type GTr(Val) mutated to TAC(Tyr)"
 PN MO9313202-A.
 PD 08-JUL-1993.
 PF 30-DEC-1992; 92WO-EP03016.
 PR 31-DEC-1991; 91IT-OMI3513.
 PA (BIOC-) BIOGINE SCLAVO SPA.
 XX PI
 XX PT Domenighini M, Hol W, Pizze M, Rappuoli R;
 DR WPI: 1993-227320/28.
 DR P-PSDB: AAR44018.
 PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PR - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli

PS Claim 3: Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051319 is a
 CC combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 235 A; 124 C; 155 G; 197 T; 0 other;

alignment_scores:
 Quality: 1178.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
 US-09-528-682-1 x AA051319 ..

Align seg 1/1 to: AA051319 from: 1 to: 711

```

1 AsnGlyAspArgLeuYrAgaAlaAspSerArgProProAspGluIleY 17
  |||
1 AATGGCGACAGATTATACCGTGTGACTAGACCCCGAGTGAATAAA 50
17 SARSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||
51 ACGTTTCCGAGCTTATGCCAGAGGT...AATGAGTACTTCGATGAG 97
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
98 GAACCTCAATGATATATTAATCTTATGATCAGCGAGAGGACCAAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
  |||
148 GCGTTTTCAGATATGATGACGAGTATGTTCCACTCTCTAGTTTGAG 197
67 gserAlHisIleuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyrT 84
  |||
198 AAGTCTCAGCTTACGACAGACGATATATATTCAGATATTCAGTTACTA 247
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
248 TATATATCGTTATAGCA.....AATATGTTTAATTCAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
  |||
289 ATTGCGCTATACACCCCTCACCCATATGACAGAGGATTCCTCGGTTAGG 338
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
  |||
339 TGGATATCCATATCTCAGATATATGATGATGATCGTTRATTTTGCTG 388
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||
389 TGATTGATGAACGATATCATCGTAAACAGGAAATATAGAACCCGTTATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgGluLeuAlaGly 167
  |||
439 AGAATATGATATATAGCTCCGCGAGAGATGTTACAGATTAGACAGTTT 488
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAlaP 184
  |||
489 CCCACCCGATACCCACAGCTTCGAGAGAAAGAACCTTGATTCATATGAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||
539 CACAAGGTTGTGAGATTTCATCAAGAACATCAACAGGTGATCTTGTAAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrIleuArgGluTyrGlnSerLy 217
  |||
589 GAGAGAGACCCAGAACTGAGACAAATATATCTCAGGAAATATCAATCAA 638
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
  |||
639 AGTTAAAGGACAGATATTTTCAGACTTCAGCTCAGAGCTTGACATATATA 688
234 snArgIleArgAspGluLeu 240

```

|||||
 689 ACAGAAATTCGAGATGAAATTA 708

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051315

seq_documentation_block:
 ID AA051315 standard; DNA; 711 BP.

AC AA051315;

XX 08-DEC-1993 (first entry)

DE Encodes Glu-53 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

FT Key Location/Qualifiers

FT misc_difference 153..155

FT /*tag= a

FT /note= "wild-type GTC(Val) mutated to GAA(Glu)"

XX WO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-0M13513.

XX (BIOC-) BIOGINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX WPI, 1993-227320/28.

XX P-PSDB; AAR38731.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

CC toxin (LT-A) of a strain of E.coli known to affect humans (Tamaoto

CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to

CC site-directed mutagenesis. Certain mutations were found to reduce

CC toxicity (see AA051314-051326). The invention relates to

CC immunogenic, detoxified LT-A proteins and their use in vaccines to

CC protect against enterotoxigenic E.coli. Sequence AA051315 is a

CC combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino

CC acid numbering is based on the cholera toxin A subunit sequence).

SO Sequence 711 BP; 236 A; 122 C; 156 G; 197 T; 0 other;

alignment_scores:
 Quality: 1177.00 Length: 240
 Ratio: 5.185 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
 US-09-528-682-1 x AA051315 ..

Align seg 1/1 to: AA051315 from: 1 to: 711

```

1 AsnGlyAspArgLeuYrAgaAlaAspSerArgProProAspGluIleY 17
  |||
1 AATGGCGACAGATTATACCGTGTGACTAGACCCCGAGTGAATAAA 50

```



```

17 sargserglyleumetproargglyhlsasnlglytyrphespargg 34
   |||||
51 ACgTTCCGGAGCTTATGCCCAAGST...AATGACTACTTCGATGAG 97
34 lYthrGlnMetAsnIleasnLeuTyraSPHsalaargglythrGlnthr 50
   |||||
98 GAACtCAAAATGAATATTAATCTTATGATCAGCGAGAGAACCAAAACC 147
51 GlypheValargTyraSPaSPglyTyraValSerThrSerLeuSerLeuAr 67
   |||||
148 GGCtTTGAAAGATATGATGACGAGTATGTTCCCTCTCTTAACTTGGAG 197
67 gSerAlaHisIleuAlaGlyInsSerIleLeuSerGlyTyraSerThrTyrt 84
   |||||
198 AAGtGCTCACTTAGCAGACAGATATATATATGATGATTCACCTTACTA 247
84 yrlleTyraIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   |||||
248 TATATATCGTTATAGCA.....AATATGTTTAAATGTTAATGATGTA 288
101 leuGlyValIlySerProHisProTyrgluGlnIuValSerAlaLeuG1 117
   |||||
289 ATTAAGCTATACAGCCCTCACCATATGAACAGAGCTTCTCGGTAGG 338
117 yGlyIleProTyraSerGlnIleTyrglyTyraTyraValAsnPhedgly 134
   |||||
339 TCGAATACCTATTTCTCAGATATATGATGATGATCGTGTATATTTGGTG 388
134 alIleasPgluArgLeuHisArgAsnArgGluTyraTARasPARGTyraTyr 150
   |||||
389 TGATGTATGAACGATTCATCGTACAGGGAATATAGAACCGGTATATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyraTArgleuAlaGlyPh 167
   |||||
439 AGAATGTGAATATAGCTCCGCGAGAGATGTTACAGATTAGCAGAGTTT 488
167 eProProAsnHisGlnIleAlaThrPARGgluGluProThrIleHisIleAlaP 184
   |||||
489 CCCACCCGATACACCACTTGAGAGAGAAACCCGTGATTCATCATGACAC 538
184 roGlnIlyCySGlyAsnSerSerArgThrIleThrGlyAspThrCyAsn 200
   |||||
539 CACAGAGTGTGGAGATTTCATCAAGAACAAATCACAGGTGATCTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyraLeuArgGluTyrglnSerly 217
   |||||
589 GAGAGAGACCCAGATCTGACACAAATATATCTCAGGGAATATCAATCAA 638
217 sValIyAsrGlnIlePheSerAspTyrglnSerGluValAspIleTyra 234
   |||||
639 AGTTAAAGGAGAGATATTTTCAGACTATACATCAGTCAGAGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
   |||||
689 ACAGAAATTCGGAGTGAATTA 708
seq_name: /STDSS/gcsgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51318
seq_documentation_block:
ID AAQ51318 standard; DNA; 711 BP.
XX
AC AAQ51318;
XX
DE 08-DEC-1993 (first entry)
XX
KM enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KM protome A; site-directed mutagenesis; reduced toxicity; ss.
XX ADP-ribosyltransferase activity; ss.
OS Escherichia coli.

```

```

XX
FH Key Location/Qualifiers
FT msc_difference 277..279
FT /tag= a
FT /note= "wild-type G7T(Val) mutated to AAG(Lys)"
XX
MO9313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91IT-OMI3513.
XX
PA (BIOC-) BIOCINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Pizza M, Rappuoli R;
XX
DR WPI; 1993-227320/28.
XX
DR P-PSDB; AAR44017.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
PS Claim 3; Fig 2 and Page 46; 60pp; English.
XX
CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AAQ51314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AAQ51318 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 BP; 236 A; 123 C; 156 G; 196 T; 0 other;

```

alignment_scores:

Quality:	1177.00	Length:	240
Ratio:	5.185	Gaps:	2
Percent Similarity:	94.583	Percent Identity:	92.917

alignment_block:

US-09-528-682-1 x AAQ51318 ..

Align seg 1/1 to: AAQ51318 from: 1 to: 711

```

1 AsnGlyaspArgLeuTyraGAlaAspSerArgProProaspGluIlely 17
   |||||
1 AATGCGACAGATTAATACCGTGTGACTCTGAGACCCCAAGATGAATAAA 50
17 sargserglyleumetproargglyhlsasnlglytyrphespargg 34
   |||||
51 ACgTTCCGGAGCTTATGCCCAAGST...AATGACTACTTCGATGAG 97
34 lYthrGlnMetAsnIleasnLeuTyraSPHsalaargglythrGlnthr 50
   |||||
98 GAACtCAAAATGAATATTAATCTTATGATCAGCGAGAGAACCAAAACC 147
51 GlypheValargTyraSPaSPglyTyraValSerThrSerLeuSerLeuAr 67
   |||||
148 GGCtTTGCAAGATATGATGACGAGTATGTTCCCTCTCTTAACTTGGAG 197
67 gSerAlaHisIleuAlaGlyInsSerIleLeuSerGlyTyraSerThrTyrt 84
   |||||
198 AAGtGCTCACTTAGCAGACAGATATATATATGATGATTCACCTTACTA 247
84 yrlleTyraIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   |||||

```


DR P-PSDB; AAR38730.

xx Immunogenic detoxified mutant cholera toxin and heat labile toxin
xx PT - useful as vaccines against infection by *Vibrio cholerae* and
xx enterotoxin producing *Escherichia coli*

PS Claim 3; Fig 2 and Page 46; 60pp; English.

xx

CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of *E. coli* known to affect humans (Yamamoto
CC et al., J. Biol. Chem., 259, 5037-5044 - see AA042768) was subjected
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic *E. coli*. Sequence AA051314 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence)

xx
xx
SQ Sequence 711 BP; 235 A; 122 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 1176.00 Length: 240
Ratio: 5.181 Gaps: 2
Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
US-09-528-682-1 x AA051314 ..

Align seg 1/1 to: AA051314 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleLeu 17
|||||
1 AATGGCGACAGATTATACCGCTGACTAGACCCCGAGTGAATATAA 50
17 sArgSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
51 ACGTTCCGGAGTCTTATGCCCAAGGT...AATGAGTCTTGATGATGAG 97
34 IYThGlMetAsnIleAsnLeuTyrSphHisAlaArgGlyThrGlnThr 50
98 GAACCAATGATATATTATCTTATGATCATCCGAGAGAACAAACC 147
51 GlyPheValArgTyrAspArgIleTyrValSerThrSerLeuSerLeuArg 67
|||||
148 GCGTTGTAGATATGAGACGGATATGTTCCACTTCTTAGTTGAG 197
67 gSerLanHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
198 AAGTGTCTACTTAGCAGACGACAGTATATATTATGAGATATCTTACTA 247
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
248 TATATATGCTTATAGCA.....AATATGTTTAAATGTTATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
289 ATTCGCGATTACAGCCCTCACCCATATGAACAGAGAGGTTCTCGCTTAA 338
117 yGluIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||
339 TGGATACCATATTTCTCGATATATGATATGATGCTGATTAATTTTGG 388
389 TGAATGATGAACGATATACATGCTAACAGGGAATATGAGACGCGTATTAC 438
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
439 AGAAATCTGAATATATGCTCCGCGACAGAGATGTTACGATATCCAGTTT 488
167 eProProAsnHisGlnAlaTyrArgGluGluTyrProTyrIleHisAla 184


```

|||||
489 CCCACCGGATCACCAGCTTGAGAGAGAACCTGGATTCATCATGCAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAAGGTGTGGAGATTCATCAAGAACATCAACAGCTGACTTGTGAAT 588
201 GlnGluThrGlnAsnLeuSerThrIleTyrlleuArgGluThrGlnSerly 217
589 GAGGAGACCCAGATCTGAGCACAATATATCTCAGGAAATATCAATCAA 638
217 sVallyArgGlnIlePheSerAspTyrlGlnSerGluValAspIleTyra 234
639 AGTTAAGAGGACAGATATTTTCAGACATCATCAGTCAAGGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /SID55/gcgbeta/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51320

seq_documentation_block:

ID AAQ51320 standard: DNA; 711 BP.

AC AAQ51320;

DT 08-DEC-1993 (first entry)

DE Encodes Gln-107 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KM protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 307..309 /tag="a

FT /note="Wild-type CAC(His) mutated to GAG(Glu)"

XX WO9313202-A.

XX PD 08-JUL-1993.

XX PR 31-DEC-1991; 91IT-0M13513.

XX PA (BIOC-) BIOCINE SCLAVO SPA.

XX PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX WPI; 1993-227320/28.

XX DR P-PSDB; AAR44019.

XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PS enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

The wild-type sequence coding for the A subunit of the heat labile toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AAQ51314-051326). The invention relates to immunogenic, detoxified LT-A proteins and their use in vaccines to protect against enterotoxigenic E.coli. Sequence AAQ51320 is a combination of the wild-type coding sequence and the mutagenic primer sequence used to introduce the preferred mutation. (Amino acid numbering is based on the cholera toxin A subunit sequence).

sq Sequence 711 BP; 234 A; 121 C; 158 G; 198 T; 0 other;

alignment_scores:

Quality: 1175.00 Length: 240
Ratio: 5.154 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AAQ51320 ..

Align seg 1/1 to: AAQ51320 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyraGlnAlaAspSerArgProProAspGluIlely 17
1 AATGGCGACAGATTATACCGTGTGACTCTAGACCCCAAGATGAAATAAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrlPheAspArg 34
51 ACCTTCCGGAGCTCTTATGCCAGAGGT...ATGAGTACTTGCATATAG 97
34 lYthrGlnMetAsnIleAsnLeuTyrlAspHisAlaArgGlyThrGlnThr 50
98 GAACCTCAATGAATATATATCTTATGATCAGCGAGAGGACAAACACC 147
51 GlyPheValArgTyrlAspArgGlyTyrlValSerThrSerleuSerleu 67
148 GGCTTTGTGAGATATGATGAGGATATGTTTCCACCTCTGTAGTTTGA 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrlSerThrTyrl 84
198 AAGTGCCTATAGCAGACAGATATATATATATATATATATATATATAT 247
84 yrlleTyrlValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATATAGCA.....AATATGTTTATGTATATGATGAT 288
101 LeuGlyValTyrlSerProHisProTyrlGlnGlnValSerAlaLeuG 117
289 AATAGCGTATACAGCCCTGAGCCATATGAAACAGAGGTTTCTGGGTAG 338
117 yGlyIleProTyrlSerGlnIleTyrlGlyTyrlTyrlArgValAsnPheGly 134
339 TGGATATCCATATTTCTAGATATATATATATATATATATATATATAT 388
134 allleAspGluArgLeuHisArgAsnArgGlyTyrlArgAspArgTyrl 150
389 TGAATTGATGAACGATTACATCGTAAACAGGAAATATAGACACCGTAT 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrlArgLeuAlaGly 167
439 AGAAATCTGAATATAGCTCCGGCAGAGAGGTTTACGATTATAGCGTTT 488
167 eProProAspHisGlnAlaThrArgGlnGluIleProTyrlIleHisHisAla 184
489 CCCACCGGATCACCAGCTTGAGAGAGAACCTGGATTCATCATGCAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTGTGAGATTTCATCAAGAACATCAACAGCTGACTTGTGAAT 588
201 GlnGluThrGlnAsnLeuSerThrIleTyrlleuArgGluThrGlnSerly 217
589 GAGGAGACCCAGATCTGAGCACAATATATCTCAGGAAATATCAATCAA 638
217 sVallyArgGlnIlePheSerAspTyrlGlnSerGluValAspIleTyra 234
639 AGTTAAGAGGACAGATATTTTCAGACATCATCAGTCAAGGTTGACATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```


seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051324

seq_documentation_block:

ID AA051324 standard; DNA: 711 BP.

AC AA051324;

DT 08-DEC-1993 (first entry)

DE Encodes Ser-106 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

FT Key Location/Qualifiers

FT misc_difference 304..306

FT /note= "wild-type CCT(Pro) mutated to AGT(Ser)"

XX MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-0M13513.

XX (BIOC-) BIOGINE SCLAVO SPA.

XX Domenghini M, Hol W, Pizza M, Rappuoli R;

XX WPI: 1993-227320/28.

XX P-PSDB: AAR44023.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PT enterotoxin producing Escherichia coli

XX PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

XX toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

XX et al. J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to

XX CC site-directed mutagenesis. Certain mutations were found to reduce

XX CC toxicity (see AA051314-Q51326). The invention relates to

XX CC immunogenic, detoxified LT-A proteins and their use in vaccines to

XX CC protect against enterotoxigenic E.coli. Sequence AA051324 is a

XX CC combination of the wild-type coding sequence and the mutagenic

XX CC primer sequence used to introduce the preferred mutation. (Amino

XX CC acid numbering is based on the cholera toxin A subunit sequence).

XX SQ Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other;

alignment_scores:

Quality: 1175.00 Length: 240

Ratio: 5.176 Gaps: 2

Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AA051324

Align seg 1/1 to: AA051324 from: 1 to: 711

```
1 AsnglyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
|||||
1 AATGGGACAGATTAACCGTGCTGACTGACCCCGAGATGAATAA 50
|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
|||||
```

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051321

seq_documentation_block:

ID AA051321 standard; DNA: 711 BP.

AC AA051321;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

OS Escherichia coli.

FT Key Location/Qualifiers

FT misc_difference 298..300

```
51 ACGTTCCGGAGCTTTATGCCAGAGGT...AATGAGTACTTCGATPAG 97
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
98 GAACCTCAATGATATTAATCTTTATGATCAGCGAGAGAAACAAACC 147
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeu 67
|||||
148 GCGTTTGTCAGATATGATGACGAGATATGTTCCACTTCCTCTAGCTT 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||
198 AAGTCTCACTTAGAGAGACATATATATATACAGATATATCTACTACT 247
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAsp 100
|||||
248 TATATATCGCTTATGCA.....AATATCTTAAATGTTATATGATGA 288
101 LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeu 117
|||||
289 ATTACGATATACAGCAGCCCATATGACAGAGGCTTTCGCGTTAGG 338
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||
339 TCGAATACCATATTTCTCAGATATATGATGATGATGATTAATTTGG 388
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||
389 TGATTCATGACGATTTACATCTGTAACAGGAGATATAGACCGGATAT 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||
439 ACAAATCTGAATATAGCTCCGCGACAGAGATGCTTACAGATTAGCAG 488
167 eProProAspHisGlnAlaThrArgGlnGlnProThrIleHisAlaP 184
|||||
489 CCCACCGGATCACCAAGCTTGAGAGAGAACACCTGGATTCATACGAC 538
184 rGlnGlnIleCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 200
|||||
539 CACAAGGTTGTGAGATTCATCAGACAAATCACAAGGTGATCTGTAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
589 GAGGAGACCCAGCAATCTGACCAATATATCTCAGGAGATATCAATCAA 638
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
639 AGTTAAAGGCGAGATATTTTTCAGACTATATCAGTACAGAGGTTGAC 688
234 sNArgIleArgAspGluLeu 240
|||||
689 ACAGAAATTCGGATGATTA 708
```



```

FT      /*tag= a
FT      /note= "wild-type TrAc(Tyr) mutated to AGT(ser)"
XX
XX      WO9313202-A.
XX
XX      08-JUL-1993.
XX
XX      30-DEC-1992; 92WO-EP03016.
XX
XX      31-DEC-1991; 91IT-OMI3513.
XX
XX      (BIOC-) BIOCINE SCLAVO SPA.
XX
XX      Domenighini M, Hol W, Pizze M, Rappuoli R;
XX      WPI: 1993-227320/28.
XX      P-PSDB: AAR44020.
XX
XX      Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX      - useful as vaccines against infection by vibrio cholerae and
XX      enterotoxin producing Escherichia coli
XX
XX      Claim 3; Fig 2 and Page 46; 60pp; English.
XX
XX      The wild-type sequence coding for the A subunit of the heat labile
XX      toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
XX      et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
XX      site-directed mutagenesis. Certain mutations were found to reduce
XX      CC toxicity (see AA051314-Q51326). The invention relates to
XX      CC immunogenic, detoxified LT-A proteins and their use in vaccines to
XX      CC protect against enterotoxigenic E.coli. Sequence AA051321 is a
XX      CC combination of the wild-type coding sequence and the mutagenic
XX      CC primer sequence used to introduce the preferred mutation. (Amino
XX      CC acid numbering is based on the cholera toxin A subunit sequence).
XX
XX      Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other:

alignment_scores:
Quality: 1174.00      Length: 240
Ratio: 5.172          Gaps: 2
Percent Similarity: 94.583      Percent Identity: 92.917

alignment_block:
US-09-528-682-1 x AA051321 ..

Align seg 1/1 to: AA051321 from: 1 to: 711

1  AAGGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
1  AATGGCGACAGATTATACCGTGTGACTAGACCCCGAGATGCAATATAA 50
17  SATGSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51  ACGTTTCGGAGTCTATGCCAGAGST...AATGACTACTGTCATGAG 97
34  LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGln 50
98  GAACSTCAATGAATATATATCTTATGATCAGCGAGAGAACCAAAAC 147
51  G1PheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
148  GCGTTTTCAGATATGATGACGGATATGTTCCACTCTCTGACTTGGAG 197
67  gSerAlaHisLeuAlaGlyIleSerIleLeuSerGlyTyrSerThrTyr 84
198  AAGTGTCACTAGACAGACAGATATATATGATGATGATATCACTTCTA 247
84  yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
248  TATATATCGTTATAGCA.....AATATGTTTAAATGTTAAAGATGA 288
101  LeuGlyValIleTyrSerProHisProTyrGluGlnIleValSerAla 117

```

```

289  ATTAGCGTTAAAGACCCCTCACCCATATGACAGAGAGTTTCGTAGG 338
117  yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnDheGly 134
339  TGGAAATACATATTCAGATATATGATGATGATGATGATGATGATGATG 388
134  aIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
389  TGAATGATGAACGATATCATCTAGAACAGGAATATAGACCGGTAATAC 438
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
439  AGAAATCTGAATATAGCTCCGCGAGAGATGTTACAGATTAGACGTTT 488
167  eProProAspHisGlnAlaTyrArgGluIleProTyrIleHisAlaIle 184
489  CCCACCGGATCCACCAAGCTTGGAGAGAAACCCTGGATTCATCATGCAC 538
184  roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539  CACAAAGGTTGTGGAGATTCATCAAGAACAAATCACAGGATGATCTGTA 588
201  GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSerLy 217
589  GAGAGAGACCCAGATCTGACACAAATATCTCAGGAAATATCAATCAA 638
217  sValIleArgGlnIlePheSerAspTyrGlnSerGlyValAspIleTyr 234
639  AGTTAAAGGAGGAGATATTTTCAGACTATACGTGAGAGTTGACATATATA 688
234  snArgIleArgAspGluLeu 240
689  ACAGAAATTCGGGATGAATTA 708

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051323
seq_documentation_block:
ID      AA051323 standard; DNA; 711 BP.
XX
XX      AA051323;
XX
XX      08-DEC-1993 (first entry)
XX
XX      Encodes Ser-104 E.coli heat labile toxin subunit A.
XX
XX      enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
XX      protomer A; site-directed mutagenesis; reduced toxicity;
XX      ADP-ribosyltransferase activity; ss.
XX
XX      Escherichia coli.
XX
XX      Key      Location/Qualifiers
XX      misc_difference 298..300
XX      /*tag= a
XX      /note= "wild-type TrAc(Tyr) mutated to AGT(ser)"
XX
XX      WO9313202-A.
XX
XX      08-JUL-1993.
XX
XX      30-DEC-1992; 92WO-EP03016.
XX
XX      31-DEC-1991; 91IT-OMI3513.
XX
XX      (BIOC-) BIOCINE SCLAVO SPA.
XX
XX      Domenighini M, Hol W, Pizze M, Rappuoli R;
XX      WPI: 1993-227320/28.
XX      P-PSDB: AAR44022.
XX
XX      Immunogenic detoxified mutant cholera toxin and heat labile toxin

```


PT - useful as vaccines against infection by *Vibrio cholerae* and
 PT enterotoxin producing *Escherichia coli*

PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of *E. coli* known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic *E. coli*. Sequence AA051323 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

alignment_scores:

Quality: 1174.00 Length: 240
 Ratio: 5.172 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AA051323 ..

Align seg 1/1 to: AA051323 from: 1 to: 711

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProAspGluIleTy 17
1 AATGGCAGAGATTATACCGTGCCTAGACCCCAAGATGAATMAA 50
17 sArGSeGlyGlyLeuMetProArgGlyHisAsnGluTyPheAspArg 34
51 ACGTTTCGGAGCTTATGCCAGAGGT...AATGACTTCTTCATAGAG 97
34 LyThrGlnMetAsnIleAsnLeuTyRAspHisAlaArgGlyThrGlnThr 50
98 GAACCTCAATGAATATTAATCTTATGATCAGCGAGAGACAGACAAACC 147
51 GlyPheValArgTyRAspAspGlyTyValSerThrSerLeuSerLeuAr 67
148 GCGTTTTCAGATATGATGAGCGATATGTTCCACTCTTATGTTTGGAG 197
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyRSerThyTyT 84
198 AAGGCTCAGCTTAGCAGACAGATATATATTCAGATATTCACCTTACTA 247
84 yRleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATATGCA.....AATATGTTTAAATGTAATGATGTA 288
101 LeuGlyValIlySerProHisProTyGluGlnIluValSerAlaLeuG 117
289 ATTAGCGTAAAGTACCCCTCAACCAATATGAACAGAGGTTTCTGCGTAGG 338
117 yGlyIleProTySerGlnIleTyRgIyTyRPyRArgValAsnPhelGly 134
339 TGGATATACCATATTCAGATATATGATGATCGTATCGTTAAATTTTGTCG 388
134 allleAspGluArgLeuHisArgAsnArgIuTyRArgAspArgTyTyR 150
389 TGAATGTGAAACGATATCATCGTAAACAGGAATATAGAACCGGTATATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyRArgLeuAlaGlyPh 167
439 AGAATGTGATATATAGCTCCGCGCAGAGATGTTTACAGATTAGACGGTTT 488
167 eProAspHisGlnAlaIleTyRArgGluGluProIleHisHisAlaP 184
489 CCCACCGGATCACCAAGCTTGGAGAGAAACCTGGATTATCATCATGCAC 538

```

```

184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAGAGTGTGGAGATTCATCAAGAAACATCACAGCTGATCTGTATAT 588
201 GLuGlnThrGlnAsnLeuSerThrIleTyRLeuArgGluTyGlnSerTy 217
589 GAGAGAACCCAGATATCTGACACAAATATCTCAGGGAATATCAACAA 638
217 sValIysArgGlnIlePheSerAspTyRInserGluValAspIleTyR 234
639 AGTTAAGAGGAGATATTTTCAGACTATTCAGTCAGAGGTTGACATATATA 688
234 sNArgIleArgAspGluLeu 240
689 ACAGAAATTCGGAGTGAATTA 708

```

seq_name: /SIDS5/gcdata/geneseq/geneseqn_emb1/NA1993.DAT:AA051322

seq_documentation_block:

ID AA051322 standard; DNA; 711 BP.

AC AA051322;

DT 08-DEC-1993 (first entry)

DE Encodes Asp-104 *E. coli* heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS *Escherichia coli*.

FH Key Location/Qualifiers

FT misc_difference 298..300 /tag= a

FT /note= "wild-type TAC(Tyr) mutated to GAT(asp)"

PN WO9313202-A.

PD 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-0M13513.

PA (BIOC-) BIOCINE SCLAWO SPA.

PI Domenighini M, Hol W, Pizza M, Rappuoli R;

DR WPI; 1993-227320/28.

DR P-PSDB; AAR44021.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by *Vibrio cholerae* and

PT enterotoxin producing *Escherichia coli*

PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of *E. coli* known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic *E. coli*. Sequence AA051322 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;


```

84 yrllefyvalillealthrallaprosanmetpheasnvalasnaspval 100
|||||
251 ATATATATGTTATAGCCACATGACCCACACATTTTAACCTTAATGATGTA 300
|||||
101 leuglyvaltyrserprohisprotyrglugingluvalseralaleucl 117
|||||
301 TTAGGGGACATACGTCCTCATCCAGATGAAACAAGATTTCGCTTAG 350
|||||
117 yglyllepertyrserglniletyrglytrpyrargvalasnphneglyv 134
|||||
351 TGGGATTCCTACTCTCCAAATATATGATGATGATGATTCATTCTTTGGGG 400
|||||
134 allleaspcgluarlgleunhisargasnargglutyrargaspargtyrtyr 150
|||||
401 TGCCTGATGACAAATTAATCATCTGTAATAGGGCTACAGAGATGATATATAC 450
|||||
151 Argasnleuasnillealaproalagluaspjlytyrarglualaglyph 167
|||||
451 AGTAACTAGATATATGCTCCACAGCAGATGCTTATMGATTTGGCAGGTTT 500
|||||
167 eProProAspHisGlnAlaIleTPArgGluGluProTrpIleHisAlaIle 184
|||||
501 CCTCCGGGACATAGAGCTTGGAGGGAAGCCGCTGATTCATCATATGCAC 550
|||||
184 roglnglcyaglyvasnserargthrilethrghlyasphrcysasn 200
|||||
551 CCGCGGGTGTGGAAATGCTCCAAAGATCATGATCATATACCTTGCAGT 600
|||||
201 GluGluTrpGlnAsnLeuSerThrIleTyrlleuArgGluTyrglnSerly 217
|||||
601 GAAAAAACCCAAAGCTAGGTAAATTCCTTGACAAATACCAATCTAA 650
|||||
217 sValysargGlnIlePheSerAspTyrglnSerGluValaspIleTyra 234
|||||
651 AGTTAAAGACAAATATTTTTCAGGCTATCAATCTGATTTGATACACATA 700
|||||
234 snArgIleArgaspGluLeu 240
|||||
701 ATGAAATTAAGATGAATTA 720
|||||
seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA51157
seq_documentation_block:
ID AAA51157 standard; cDNA; 777 BP.
XX AAA51157;
AC
XX
XX
XX 26-SEP-2000 (first entry)
DE
XX Plant-optimized V. cholerae toxin A subunit coding sequence.
XX
XX Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
XX
XX adjuvant; anti-bacterial; ss.
XX
XX Vibrio cholerae.
OS
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..777 /*tag= a
XX FT /product= cholera_toxin_subunit_A
XX
XX MO200037609-A2.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99MO-US30747.
XX
XX 22-DEC-1998; 98US-0113507.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
XX
XX (MASO/) MASON H S.
XX
XX (ARNT/) ARNTZEN C J.

```

```

XX
PI Mason HS, Arntzen CJ;
XX
XX WPI: 2000-442653/38.
DR
DR P-PSDB: AAY96653.
XX
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
PS Disclosure; Fig 4A; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a synthetic Vibrio cholerae
CC cholera toxin (CT) A subunit (CT-A). The sequence has no cryptic signal
CC sequence and no CG (potential methylation sites) sequences. Novel
CC polynucleotides encode a mutant Escherichia coli heat-labile toxin (LT)
CC A subunit A (LT-A) polypeptide or a mutant V. cholerae cholera toxin (CT)
CC A subunit (CT-A) polypeptide, which have reduced enzyme activity as
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
CC of the codons is altered to a plant preferred codon. The polynucleotide
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
CC a CT B subunit (CT-B). The polynucleotides are useful for the
CC transformation of plant cells for the production of transgenic plants to
CC produce edible vaccines, especially oral vaccines in transgenic plants
CC for the prophylactic or therapeutic treatment against E. coli or
CC V. cholerae. The mutant polypeptides are also useful as adjuvants.
XX
XX
SQ Sequence 777 BP; 209 A; 174 C; 191 G; 203 T; 0 other;
XX
XX
alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667
XX
XX
alignment_block:
US-09-528-682-1 x AAA51157 ..
XX
XX
Align seg 1/1 to: AAA51157 from: 1 to: 777
XX
1 AsnGlyAspArgLeuTyraArgAlaAspSerArgProProAspGluIleTy 17
|||
55 AATGATGACAAAGCTCATAGGGCAGACTCAAGACCTCCGATGATGATCA 104
|||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyryPheAspArg 34
|||
105 GCAATCAGGTGGTCTTATGCCAAGGGGACAAATCTGATCTTGAACAGG 154
|||
34 JyThcGlnMetAsnIleAsnLeuTyraAspHisAlaArgGlyThrGlnThr 50
|||
155 GTACTGAGATGAAACATCAACCTTATGACCATGCAAGGGAACTCAACT 204
|||
51 GlyPheValArgTyraAspAspGlyTyraValSerThrSerLeuSerLeu 67
|||
205 GGATTTGTGAGGCATGATGATGATGATGATGATGATGATGATGATGATG 254
|||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrySerThrTyT 84
|||
255 GTCTGCCCACTTGTGGGTCAAACTATCTCTCTGCTGCTACTACTACT 304
|||
84 yrllefyvalillealthrallaprosanmetpheasnvalasnaspval 100
|||||
305 ACATCTATGATGATGCTGCACTGCCAACATATCTCAATGTGAATGATG 354
|||||
101 leuglyvaltyrserprohisprotyrglugingluvalseralaleucl 117
|||||
355 TTGGGAGCATACAGCCCTCACAGATGAGCAAGAGCTGCTGCTTTGGG 404
|||||
117 yglyllepertyrserglniletyrglytrpyrargvalasnphneglyv 134
|||||
405 TGGAAATCCATACCTCCAAATATATGATGATGATGATGATGATGATGATG 454
|||||
134 allleaspcgluarlgleunhisargasnargglutyrargaspargtyrtyr 150
|||||

```



```

445 1GCTTGAATAGGCACTCCATATAGGATAGGAGGCTACAGGAGATAGGACTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPac 167
505 AGCAACTTGGACATTCGCTCCACACAGACAGATGGTTATGGATTGGCAGGTTT 554
167 eProProAspHisGlnAlaTyrPargGluGluProTyrPleHisIleAlaP 184
555 CCCTCAGACAGCATAGGCGCTTGAGAGGAGAGACCTTGGATTTCACCATGCAC 604
184 roGlnGlyCYcSGLyAsnSerSerArgThrIleThrGlyAspThrCYcAsn 200
605 CACCAAGGTTGTGGAATCTCCCAAGTCACAGCATGACCAACACACTGTGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIly 217
655 GAAAGAACCCAAATCTTGGGTGTGAAGTTCCTGATGAGTACCAATCTAA 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
705 GGTGAAGAGGCAAAATCTTCTCAGCGCTACCACATTCACATTCAGCACCA 754
234 snArgIleArgAspGluLeu 240
755 ATGATGATCAAGATGATGACTC 774

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:AAAN30169
seq_documentation_block:
ID  AAAN30169 standard; DNA: 1148 BP.
XX
XX  AAAN30169;
XX
XX  25-APR-1992 (first entry)
XX
XX  Sequence of PR1F 18014 encoding sub-unit A and B of cholera toxin.
XX
XX  Vaccine; cholera; subunit A; subunit B; ss.
XX
XX  Vibrio cholerae.
XX
XX  Key Location/Qualifiers
XX  F1 1..777
XX  F1 misc_feature /*tag= a
XX  F1 /label= F1
XX  F1 /note= "claimed fragment"
XX  F1 misc_feature 774..1148
XX  F1 /*tag= b
XX  F1 /label= F2
XX  F1 /note= "claimed fragment"
XX
XX  EP95452-A.
XX
XX  30-NOV-1983.
XX
XX  23-MAY-1983; 83EP-0091416.
XX
XX  24-MAY-1982; 82US-0381083.
XX
XX  (SMIK ) SMITH KLINE-RIT.
XX
XX  Harford N, Dewilde M;
XX
XX  WPI: 1983-834665/49.
XX
XX  DNA sequence coding for cholera toxin sub units - for
XX  PT transforming host cells, useful in vaccine prodn.
XX
XX  Claim 2; Page 27-28; 46pp; French.
XX
XX  The inventors claim a sequence conty. at least one fragment coding
XX  CC for all or part of sub-units A and B of cholera toxin. The entire

```

[illegible]


```
|||||
705 AGTTAAAGACAAATTTTTCAGGCTATCAATCTGATATGTACACATA 754
234 snArg1leArgAspGluLeu 240
755 ATAGAAATTAAGATGAATTA 774
seq_name: /SIS5/gcgdata/geneseq/geneseqn-emb1/AA2001A.DAT:AA167270
seq_documentation_block:
ID AA167270 standard; DNA: 2056 BP.
XX AA167270:
XX
XX 11-FEB-2002 (first entry)
XX
XX Cholera toxin A and B subunits encoding DNA sequence.
XX
XX Enterotoxin; mutation; chimeric; immunostimulant; immune response;
XX cholera toxin; ds.
XX
XX Vibrio cholerae.
XX
XX Key Location/Qualifiers
FH 1..550
FT 5'UTR /*tag- a
FT CDS 551..1327
FT /*tag- b
FT /*product- "Cholera toxin A subunit"
FT /*note- "AAG65991"
FT CDS 1324..1698
FT /*tag- c
FT /*product- "Cholera toxin B subunit"
FT /*note- "AAG65992"
FT /*product- "B chain"
FT 3'UTR 1699..2056
FT /*tag- d
XX
XX NO200170257-A1.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001MO-US08582.
XX
XX 17-MAR-2000; 2000US-190058P.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Mcghee J, Kiyono H, Takeda Y, Ohmura M, Yamamoto S;
XX
XX WPI: 2001-648368/74.
XX
XX P-PSDB: AAG65991, AAG65992.
XX
XX Chimeric molecule useful as adjuvant for cell-mediated/humoral immunity
XX comprises first mutated A subunit of first enterotoxin and second
XX non-mutated subunit from second enterotoxin different from natural
XX enterotoxin -
XX
XX Disclosure: Fig 4; 23pp; English.
XX
XX The invention provides a chimeric molecule that comprises a first subunit
XX which is mutated A subunit of first enterotoxin and a second non-mutated
XX subunit from a second enterotoxin which is different from the natural
XX enterotoxin which has been mutated to provide A subunit. A composition of
XX matter comprising the chimeric molecule in a pharmaceutically acceptable
XX carrier is useful for obtaining enhanced immune response of an organism
XX to an antigen by administering it to the antigen. The chimeric molecule
XX is useful to provide specific immune response to a particular
XX enterotoxin, as adjuvants for use with unrelated vaccines, and to
XX customize adjuvants to direct production of cell-mediated or humoral
XX immune responses. It is also useful as mucosal adjuvant for cell-mediated
XX or humoral immunity. The present sequence represents the DNA sequence of
XX Cholera toxin A and B subunits.
```

```
XX
SQ Sequence 2056 BP; 603 A; 327 C; 479 G; 647 T; 0 other:
Alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667
alignment_block:
US-09-528-682-1 x AA167270 ..
Align seg 1/1 to: AA167270 from: 1 to: 2056
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluLeu 17
|||
605 AATGATGATAGTAAGTATATCGGCAGATCTTACACCTCGATGAATAA 654
17 SARSerGlyGlyLeuMetProArgLysAsnGlyTyrPheAspArg 34
|:::|
655 GCAGTCAGGTGGCTTATGTCAGAGAGACAGAGTGAAGTACCGAG 704
34 LThrGlnMetAsnLeuLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
705 GTACTCAATGAAATATCAACCTTATGATCATGCAAGAGAACTGAGAG 754
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
755 GGATTTGTTAGGCACGATGATGATGTTTCACCTCAATTAAGTTTGGAG 804
67 gSerHisHisLeuAlaGlyGlnSerLleLeuSerGlyTyrSerThrTyr 84
|||||
805 AAGTGGCCACTTATGTCGTAACCTATATGTCGTCCTTACTTATTT 854
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
855 ATATATATGTTATAGCAGCTGACCCACCAATGTTTAAGCTTAATGATGA 904
101 LeuGlyValTyrSerProHisProTyrGlnGlnValSerAlaLeuGln 117
|||||
905 TTAGGGGCATACAGTCCATCCAGATGAACAAGATTTGCTTACGTTAGG 954
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyVal 134
|||||
955 TGGGATTCATCTACCTCCANATATATGATGATGATGATGATTTTGGGG 1004
134 AlIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
|||||
1005 TCGTTGATGACAAATATACATCGTAATAGGGCGCTACAGATATGATTTTAC 1054
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
|||||
1055 AGTAACTTAGATATGCTCCACAGCAGATGGTATGATTTGGCAGCTTT 1104
167 eProProAspHisGlnAlaTyrArgGlnGluProTyrIleHisHisAlaP 184
|||||
1105 CCGTCGCGACATPAGAGCTTGGAGGAGAGCCGCTGATTCATCATGACAC 1154
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||
1155 CCGCGGGTGTGGGAATGCTCCAAAGATCATCGATGATGATTAATCTCGAT 1204
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSerIy 217
|||||
1205 GAAAAAACCAAGTCTAGCTTAATAATTCCTTGACGAATACCAATCTTA 1254
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAlaIleTyrA 234
|||||
1255 AGTTAAAGACAAATATTTTTCAGGCTATCAATCTGATATGTACACATA 1304
234 snArg1leArgAspGluLeu 240
|||||
1305 ATAGAAATTAAGATGAATTA 1324
```



```

XX XX Vibrio cholerae.
XX XX
XX XX US5965406-A.
XX XX
XX XX 12-OCT-1999.
XX XX
XX XX 07-JUN-1995; 95US-0488246.
XX XX
XX XX 04-AUG-1993; 93US-0102387.
XX XX 07-JUN-1984; 84US-0618199.
XX XX 27-JUN-1991; 91US-0722484.
XX XX 25-APR-1985; 85US-0726808.
XX XX 07-JUN-1985; 85US-0742534.
XX XX 22-DEC-1989; 89US-0456095.
XX XX 14-JUN-1990; 90US-0538276.
XX XX
XX XX (SERA-) SERAGEN INC.
XX XX
XX XX Murphy JR;
XX XX
XX XX WPI; 1999-632431/54.
XX XX P-PSDB; AAY55890.
XX XX
XX XX Recombinant DNA molecule encoding a three part hybrid protein used in
XX XX the treatment of Aids and genetic deficiency diseases -
XX XX
XX XX Example 2; Fig 5; 31pp; English.
XX XX
XX XX The invention relates to a recombinant DNA molecule encoding a hybrid
XX XX protein comprising three parts: (a) the first part comprises a portion
XX XX of the binding domain of a cell-binding polypeptide ligand allowing the
XX XX hybrid protein to bind to an animal cell; (b) the second part comprises
XX XX a portion of a translocation domain of a naturally occurring protein
XX XX selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
XX XX toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
XX XX toxin, which translocate the third part of the across the cytoplasmic
XX XX membrane into the cytosol of the cell; and (c) the third part comprises
XX XX a polypeptide entity to be introduced into the cell, which is non-native
XX XX to the naturally occurring protein of (b). This sequence represents the
XX XX Vibrio cholerae toxin A1 fragment gene for use in generating the hybrid
XX XX of the invention. The hybrid molecule enables the direction of
XX XX appropriate therapy to affected cells, allowing them to function properly
XX XX and alleviate or cure the disease. The hybrid is especially used in
XX XX treating genetic deficiency diseases, by delivering to affected cells
XX XX an enzyme supplying the missing function, to supplementing cellular
XX XX levels of a particular enzyme or a scarce precursor or cofactor, to
XX XX directing toxins or other poisons to destroy particular cells (such as
XX XX adipocytes, cancer cell, or virus infected-cells), to counteracting viral
XX XX infections such as HIV, by introducing appropriate antibodies to viral
XX XX proteins. It is also involved in the process of getting non-therapeutic
XX XX substances such as detectable labels into cells.
XX XX
XX XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:
XX XX
XX XX alignment_scores:
XX XX      Quality: 1085.00      Length: 240
XX XX      Ratio: 4.801      Gaps: 0
XX XX Percent Similarity: 94.167      Percent Identity: 81.250
XX XX
XX XX alignment_block:
XX XX US-09-528-682-1 x AAZ30661 ..
XX XX
XX XX Align seg 1/1 to: AAZ30661 from: 1 to: 780
XX XX
XX XX 1 AAGGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIlely 17
XX XX      ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
XX XX 55 AATGAGCATGATTAAGTTATCGGCGAGATTCTAGACCTCGTGAATVAAA 104
XX XX
XX XX 17 sATGSeGlyGlyLeuMetProArgGlyHisAsnGlnTyrPheAspArg 34
XX XX      |||:|||||:|||||:|||||:|||||:|||||:|||||
XX XX 105 GCGATCGGGTGTCTATATGCCAAGAGGACAGAGTCAGTACTTTTACCGAG 154

```

34	lythrglnmetasnllleasnleuylrnsphnsialaargglythrglnthr	50
155	gltactcaaaatgaatgatcacctttatgatcwtgcgaaggaacacacagc	204
51	glpphenalargtyrarspaspglytyrvalserthrsleuslerleuar	67
205	gcatrttttaggcacacgatgatgatgtttccacstccatattagttag	254
67	gseralnlisleualaglyglnserlleleuserglytyrserthryt	84
255	aagggccaccattagtggtcacaactaatattgtcgtgcatctactat	304
84	ytlllelyvalllleallthrallprosnmetpreasnvalasaspval	100
305	atataatattgttatagccactgcaccaccaacatgtttaacgtttaagat	354
101	leugllyalutyserprohlsprotuglulglnluvalseralaleugl	117
355	tttagggcgatracacgtcctcatccagatgacacagactttgcttagg	404
117	yglylleprotyrserglnlletyrglytyrtyrtyrargvalasnphel	134
405	tgggattccactactcccaaatatgatgatggtatcgatcgattcat	454
134	allleasergluatrgleuhsargasnargluityrtyrtyrtyrtyr	150
455	tgccttgatggaacaaattatcatcgtatgatggcgctacagagatga	504
151	argasnleuasnlllealproalaglaspglytyrtyrtyrtyrtyr	167
505	agtaacttagatattgctccacgacgacagatggtatgatgtggcag	554
167	eprproasphtsglnalatrpargglugluprotprllehshslalap	184
555	ccctcccgacacatgacgttgagaggaagcccggtgcatcatgcac	604
184	roglnglucyglyasnserserarphrlllethnglasphtcysasn	200
605	cggcggtgtgggaatgctccaaagatcatcgatgacgaatgacgtg	654
201	glugluthrglnasnleuserthrilletyrleuargglutyrglnser	217
655	gaataaacccaaagctagtgtaaaatttccttgacgatatccaatcca	704
217	svalysarglnllepheseraspyrglnsergluvalasplletyra	234
705	agtttaaaagcaaatatttttcagcgctatcatcgtatgatatgacac	754
234	snarglleargaspgluleu	240
755	atgaatgtaagatgaattta	774
seq_name: /sids5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AA290017		
seq_documentation_block:		
ID	AA290017	standard; DNA; 780 BP.
XX	AA290017;	
AC	05-MAY-2000	(first entry)
DT		
XX		
DE		
XX		
KM		
XX		
KM		
XX		
OS		
XX		
PN		
XX		
PD		


```

XX 08-JUL-1993.
XX
XX 30-DEC-1992; 92WO-EP03016.
XX
XX 31-DEC-1991; 91IT-0MI3513.
XX
XX (BIOC-) BIOCINE SCLAVO SPA.
XX
XX Domenighini M, Hol W, Pizsa M, Rappuoli R;
XX
XX WPI; 1993-227320/28.
XX
XX P-PSDB; AAR44027.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX - useful as vaccines against infection by Vibrio cholerae and
XX enterotoxin producing Escherichia coli
XX
XX
XX Claim 3; Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
XX toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
XX was subjected to site-directed mutagenesis. Certain mutations were
XX found to reduce toxicity (see AA051327-051334). The invention relates
XX to immunogenic, detoxified CT-A proteins and their use in vaccines
XX to protect against Vibrio cholerae. Sequence AA051328 is a
XX combination of the wild-type coding sequence and the mutagenic
XX primer sequence used to introduce the preferred mutation.
XX
XX Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

```

```

alignment_scores:
    Quality: 1084.00      Length: 240
    Ratio: 4.796          Gaps: 0
    Percent Similarity: 94.167    Percent Identity: 81.250

```

```

alignment_block:
US-09-528-682-1 x AA051328 ..

```

```

Align seg 1/1 to: AA051328 from: 1 to: 723

```

```

1 AaNGlYAspArgLeuTyrArgAlaAspSerArgProProaspGluLeu 17
111
1 AATGATGATTAAGTATATCGGCGAGATTCTAGACCTCGATGAAATTA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
111
51 GCAGTCAGGAGGCTTATGCGCAGAGACAGACTGAGTCTTGACCGAG 100
34 lYThrGlnMetAsnLeuAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGAC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
111
151 GGAATTGTTAGCGACGATGATGATGTTCTCCACCAAGATTGTTGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerLeuSerGlyTyrSerThrTyrT 84
111
201 AAGTGGCCCACTTGTGCTCAACCTATATTGTCGTGCTCATTTCACTTA 250
84 yTLeuTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCACTGCGACCCACACATGTTAACTTAATGATGA 300
101 LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG 117
111
301 TTATGGGCGAATACGCTCTCATCCAGATGAACAAGATTCTCTCTTAGC 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
351 TGGGATTCCATCTCCCAATATATGATGATGATGATTCATTTGGGG 400

```

```

134 aAlLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
111
401 TGCTTGATGACACATTATCATCGTAAATGAGGCGTACAGACATTAATTTAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
111
451 AGTAACTTATGATTTGCTCCAGCAGACAGATGTTATGATTTGACAGTTT 500
167 eProProAspHisGlnAlaThrArgGluGluProTrrPheHisAlaP 184
111
501 CCTCCGAGCATATAGACTTGGAGGAGAGAGCCGCTGATTCATCATGACAC 550
184 rGlnGlyCysGlyAsnSerSerArgThrLeuThrGlyAspThrCysAsn 200
111
551 CGCCGGGTTGTGGAAATGCTCCAAAGATCATCGATCATGATTAATCTGCGAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
111
601 GAAAAAACCCAAAGTCTAGTGTAATAATTCCTTGACGAATACCAATCTAA 650
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
111
651 AGTTAAAGACAATATTTTCAGCTATCATCTCATGATTTGATGATACACATA 700
234 sNArgIleArgAspGluLeu 240
111
701 ATGAAATTAAGATGAATTA 720

```

```

seq_name: /SID5/gcdata/geneseq/emb1/NA1993.DAT:AA051334

```

```

seq_documentation_block:

```

```

ID AA051334 standard; DNA: 723 BP.

```

```

AC AA051334;

```

```

DT 08-DEC-1993 (first entry)

```

```

DE Encodes Glu-114 cholera toxin subunit A.

```

```

XX Immunogenic detoxified CT-A; cholera vaccine; adenylylate cyclase;

```

```

XX Protomer A; site-directed mutagenesis; reduced toxicity;

```

```

XX ADP-ribosyltransferase activity; ss.

```

```

OS Vibrio cholerae.

```

```

XX Key Location/Qualifiers

```

```

XX misc_difference 340..342 /tag= a

```

```

XX FT /note= "wild-type TCT(Ser) mutated to GAA(Glu)"

```

```

XX WO9313202-A.

```

```

XX 08-JUL-1993.

```

```

XX 30-DEC-1992; 92WO-EP03016.

```

```

XX 31-DEC-1991; 91IT-0MI3513.

```

```

XX (BIOC-) BIOCINE SCLAVO SPA.

```

```

XX Domenighini M, Hol W, Pizsa M, Rappuoli R;

```

```

XX WPI; 1993-227320/28.

```

```

XX P-PSDB; AAR44033.

```

```

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

```

```

XX - useful as vaccines against infection by Vibrio cholerae and

```

```

XX enterotoxin producing Escherichia coli

```

```

XX Claim 3; Fig 2 and Page 48; 60pp; English.

```

```

XX The wild-type sequence coding for the A subunit of the cholera

```



```

|||||
151 GGAATTTGTAGCAGCATGATGATATGTTCCACCCATATGATTGAG 200
67 gserAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
201 AAGTGGCCCATGTAAGTGGTCAACATATATGCTGTGATCTTCAATTTAT 250
84 yrlleYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCATGTCACCCACATGTTTAACTGTTATGATGA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCTCATCCAGATTCACAAAGAACTTCTGCTT 350
117 yglYlleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
351 TGGGATTCATCTACTCCCAATATATGATGATGATCGATTCAATTTGGG 400
134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
401 TCGTGTATGAAACAATTACATCGTATATAGGGGCTACAGAGATATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
451 AGTAACTTATGATATGCTCCAGCAGCAGATGTTATGATTCGACAGGTT 500
167 eProProAspHisGlnAlaThrParGluGluProTyrPheHisHisAla 184
501 CCCCTCCGACATAGACCTTGAGGGAAGAGCCGTGATTCATCATGTCAC 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CGCGGGTGTGGGAATGCTCAAGATCATGATCATATATCTGGAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
601 GAAAAAACCCAAAGTCTAGGTGTAATAATCTCTGACGAATACCAATCTAA 650
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
651 AGTAAAGACAAATATTTTCAGGCTATCATCTGATATGATGATACACATA 700
234 snArgIleArgAspGluLeu 240
701 ATAGAAATTAAGATGAATTA 720

seq_name: /stids/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51329
seq_documentation_block:
ID AAQ51329 standard; DNA; 723 BP.
XX
AC AAQ51329;
XX
DT 08-DEC-1993 (first entry)
XX
DE Encodes Lys-97 cholera toxin subunit A.
XX
KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
KW protomer A; site-directed mutagenesis; reduced toxicity; ss.
XX
KW ADP-ribosyltransferase activity; ss.
XX
OS Vibrio cholerae.
XX
FH Key location/Qualifiers
FT misc_difference 289..291
FT /tag= a
FT /note= "wild-type G7T(Val) mutated to AAG(Lys)"
XX
XX MO9313202-A.
XX
XX 08-JUL-1993.
XX

```

```

PF 30-DEC-1992; 92MO-EP03016.
XX
XX 31-DEC-1991; 91IT-0MI3513.
PR
XX
XX (BIOC-) BIOCINE SCLAVO SPA.
PA
XX
XX Domenighini M, Hol W, Piazza M, Rappuoli R;
PI
XX
XX WPI: 1993-227320/28.
DR
XX
XX P-PSDB; AAR44028.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX
XX - useful as vaccines against infection by Vibrio cholerae and
XX
XX enterotoxin producing Escherichia coli
XX
XX
XX Claim 3; Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
XX
XX toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AAQ46318)
XX
XX was subjected to site-directed mutagenesis. Certain mutations were
XX
XX found to reduce toxicity (see AAQ51327-051334). The invention relates
XX
XX to immunogenic, detoxified CT-A proteins and their use in vaccines
XX
XX to protect against Vibrio cholerae. Sequence AAQ51329 is a
XX
XX combination of the wild-type coding sequence and the mutagenic
XX
XX primer sequence used to introduce the preferred mutation.
XX
XX
XX Sequence 723 BP; 229 A; 125 C; 163 G; 206 T; 0 other;
XX

```

```

alignment_scores:
Quality: 1082.00 Length: 240
Ratio: 4.809 Gaps: 0
Percent Similarity: 93.750 Percent Identity: 81.250

```

alignment_block:

US-09-528-682-1 x AAQ51329 ..

Align seg 1/1 to: AAQ51329 from: 1 to: 723

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
||| |||
1 AAGATGATATAGTATATCGGCGAGATCTAGACCTCGATGATCAATTA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
||| |||
51 GCAGTCAGGTGGTCTTATGCCAAGAGACAGAGTGAATCTTGACCGAG 100
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGCACTCAGAG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||
151 GGATTTGTTAGCGCAGATGATGATATGTTTCCACCTCAATAGTTTGAG 200
67 gserAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
201 AAGTGGCCCATGTAAGTGGTCAACATATATGCTGTGATCTTCAATTTAT 250
84 yrlleYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCATGTCACCCACATGTTTAACTGTTATGATGA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCTCATCCAGATGACACAAAGAACTTCTGCTT 350
117 yglYlleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
351 TGGGATTCATCTACTCCCAATATATGATGATGATCGATTCAATTTGGG 400
134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
||| |||

```


[illegible][illegible]

234 snArglleArgspgluLeu 240
 |||||||:|||||||
 701 ATAGATTAAAGTCAATTAA 720

seq_name: /SIDS5/gcdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051327

seq_documentation_block:

ID AA051327 standard; DNA: 723 BP.

XX AA051327:

XX 08-DEC-1993 (first entry)

DE Encodes Asp-53 cholera toxin subunit A.

KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

FT Key Location/Qualifiers

FT misc_difference 156..158

FT /*tag="a" /note="wild-type GTT(Val) mutated to GAC(Asp)"

XX MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-0MT3513.

XX (BIOC-) BIOCTINE SCLAVO SPA.

XX Domenighini M, Hol W, Piazza M, Rappuoli R;

XX WPI: 1993-227320/28.

XX P-PSDB: AAR44026.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 48; 60pp; English.

XX The wild-type sequence coding for the A subunit of the cholera
 CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
 CC was subjected to site-directed mutagenesis. Certain mutations were
 CC found to reduce toxicity (see AA051327-051334). The invention relates
 CC to immunogenic, detoxified CT-A proteins and their use in vaccines
 CC to protect against Vibrio cholerae. Sequence AA051327 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation.

XX Sequence 723 BP; 228 A; 126 C; 163 G; 206 T; 0 other;

alignment_scores:

Quality: 1081.00 Length: 240

Ratio: 4.804 Gaps: 0

Percent Similarity: 93.750 Percent Identity: 81.250

alignment_block:

US-09-528-682-1 x AA051327 ..

Align seg 1/1 to: AA051327 from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
 ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1 AATGATGATTAAGTATATCGGCAATCTAGACCTCGATGATAATAA 50

17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 51 GCAGTCAGGTGCTTATGCCAAGAGGACAGTACGTAATCTTGACCGAG 100
 34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 101 GTACTCAATTAATGAAATATCAACCTTATGATCATGCAAGAAGAACTCAGACG 150
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 151 GGATTTGACGACGACGATGATGATATGTTTCCACCTCAATTAAGTTGAG 200
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 201 AAGTCCCACTTATGCGGTGCAACATATATGCTGCTGCTTACTTATNT 250
 84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 251 ATATATATGTTATAGCCACTGCACCAACATGTTAAACGTTAATGATGA 300
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 301 TTAGGGCATACAGTCCCTATCCAGATGACAGAGAGATTTCGCTTTAGG 350
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyV 134
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 351 TGGGATTCATCTACTCCCAATATATGAGATGATGAGTTCATTTTGGGG 400
 134 aIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 401 TCCTTGATGACAAATATACATCGTAATAGGGCTACAGATATGATATATAC 450
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 451 AGTAACTTATGATATGCTCTCAGCAGCAGATGCTTATGCTTGGCAGTTT 500
 167 eProProAspHisGlnAlaTPrArgGluGluProTPrIleHisAlaAlaP 184
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 501 CCTCGGAGCATAGAGCTTGAGGGGAGAGCCGTGATTCATCATGCAC 550
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 551 CCGCGGGTGTGGGAAATGCTCCCAAGATCATGATCATGATATCTGCGAT 600
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 601 GAAAAAACCAGTCTAGGTGAATAATCTTGACGATACCAATCTTA 650
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 651 AGTTAAACAAATATTTTTCAGGCTATCATCTGATATGATACACANA 700
 234 snArglleArgspgluLeu 240
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 701 ATAGATTAAAGTCAATTAA 720

seq_name: /SIDS5/gcdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051331

seq_documentation_block:

ID AA051331 standard; DNA: 723 BP.

XX AA051331:

XX 08-DEC-1993 (first entry)

DE Encodes Asn-107 cholera toxin subunit A.

KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by *Vibrio cholerae* and
 PT enterotoxin producing *Escherichia coli*

PS Claim 3; Fig 2 and Page 48; 60pp; English.

XX The wild-type sequence coding for the A subunit of the cholera
 CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
 CC was subjected to site-directed mutagenesis. Certain mutations were
 CC found to reduce toxicity (see AA051327-Q51334). The invention relates
 CC to immunogenic, detoxified CT-A proteins and their use in vaccines
 CC to protect against *Vibrio cholerae*. Sequence AA051330 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation.

XX Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

alignment_scores:

Quality: 1080.00 Length: 240
 Ratio: 4.800 Gaps: 0
 Percent Similarity: 93.750 Percent Identity: 81.250

alignment_block:

US-09-528-682-1 x AA051330 ..

Align seg 1/1 to: AA051330 from: 1 to: 723

```

1 AsnGlyspargLeuTyArgAlaAspSerArgProProspgIuIlely 17
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 AATGATATAGTATATCGGCGAGATTCTAGACCTCTGTAATAAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyArgPheaspArg 34
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 GCATCTCAGGTGGTCTTATGCCAAGAGACAGAGAGTACTTTGACCGAG 100
34 lyThrGlnMetAsnIleAsnLeuTyArgPheAspHisAlaArgGlyThrGlnThr 50
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 GTACTCAATGAATATCAACCTTATGATCATGCAAGAGAACACAGACG 150
51 GlyPheValArgTyArgAspArgGlyTyValSerThrSerLeuSerLeuArg 67
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GGATTTGTAGGCGACAGATGATGATGTTCCACCTCAATTAAGTTAGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyArgSerThrTyArg 84
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 AAGTCCCACTTACGTGGGTCAACTATATGTCGTGATCTTACTTATAT 250
84 yIleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 ATATATATGTTATAGCCACCTGACCAACATGTTTAACGTTAATAGATGA 300
101 LeuGlyValTyArgProHisProTyArgGlnGluValSerAlaLeuGly 117
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 TTAGGGCCATACACTAGCCATCCAGATGAACAAAGATTTCTGCTTAGG 350
117 yGlyIleProTySerGlnIleTyArgGlyTyArgValAsnPheGlyVal 134
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 TGGGATTCCTACTCCCAATATATGATGATGATGATGATGATGATGATG 400
134 alIleAspGluArgLeuHisArgAsnArgGluTyArgAspArgTyArg 150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 TGCCTGTAAGAACATATACATCGTAAATAGGGGCTACAGAGATATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgLeuAlaGlyPhe 167
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 AGTAACTTAGATATGCTCCAGCGACAGATGATGATGATGATGATGATG 500
167 eProProAspHisGlnAlaArgPheArgGluGluProTyArgIleHisHisAla 184
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 CCTCCCGAGATAGAGCTTGGAGGGAAGGCCCTGGATTCATCATCATCAC 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200

```

seq_name: /SIDS5/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:AA051160

seq_documentation_block:

```

ID AA051160 standard; cDNA; 777 BP.
XX
XX AA051160;
AC
XX 26-SEP-2000 (first entry)
XX
XX Plant-optimized mutant V. cholerae toxin A subunit G192 coding sequence.
XX
XX Heat-labile toxin: CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
XX adjuvant; anti-bacterial; R192G; SS.
XX
XX Vibrio cholerae.
XX
XX Synthetic.
XX
XX Key location/Qualifiers
XX CDS 1..777
XX FT /*tag= a
XX FT /product= mutant_cholera_toxin_subunit_A
XX
XX WO200037609-A2.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US30747.
XX
XX 22-DEC-1998; 98US-0113507.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
XX (MASO/) MASON H S.
XX (ARNT/) ARNTZEN C J.
XX
XX Mason HS, Arntzen CJ;
XX
XX MPI: 2000-442653/38.
XX P-PSDB: AAY96656.
XX
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
XX transformation of plant cells, useful in immunogenic compositions to
XX elicit immune responses in animals
XX
XX Claim 15; Fig 4D; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a mutant synthetic
XX Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a R192G
XX substitution. The sequence has no cryptic signal sequence and no CG
XX (potential methylation sites) sequences. Novel polynucleotides encode a
XX mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
XX polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
XX polypeptide, which have reduced enzyme activity as compared to the
XX wild-type LT-A or CT-A polypeptide and where at least one of the codons
XX is altered to a plant preferred codon. The polynucleotide further
XX comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
XX subunit (CT-B). The polynucleotides are useful for the transformation of
XX plant cells for the production of transgenic plants to produce edible
XX vaccines, especially oral vaccines in transgenic plants for the

```


CC catalytic RNAs in animal cells or tissues. The bacterial blebs

CC containing the eukaryotic expression cassette can also be used to treat
 CC animal cells cultured in vitro. The method can be used in gene therapy.
 CC The present sequence represents the cholera toxin A1 subunit sequence,
 CC which is used in an example from the present invention for the production
 CC of non-pyrogenic blebs containing a DNA vaccine.

Sequence 7586 BP; 1619 A; 1747 C; 1733 G; 1657 T; 830 other;

alignment_scores:

Quality:	929.00	Length:	191
Ratio:	5.049	Gaps:	0
Percent Similarity:	96.335	Percent Identity:	86.911

alignment_block:

US-09-528-682-1 x ABA92644 ..

Align seg 1/1 to: ABA92644 from: 1 to: 7586

```

3 AsparGleuTyrrArgAlaAspSerArgProPoaSgLuIleLysArgse 19
|||||
9 GATAGATTATCGGCAGATCTAGACCTCGATGAATTAAGCAGCTC 58
19 rgllyGlyLeuMetProArgGlyHisAnGluTyrrPheAspArgGlyThrc 36
|||||
59 AGGTGGCTTATGCCAAGAGCAGAGTAGTACTTTGACCGAGTACTC 108
36 lMetSnlleAsnLeuTyrrAspHisAlaArgGlyThrcGlnThrcGlyPhe 52
|||||
109 AATGATATACCACTTTATGATCATGCAAGAGAACTCAGACGGGATT 158
53 ValArgTyrrAspAspGlyTyrrValSerThrSerLeuSerLeuArgSerAl 69
|||||
159 GTTAGGCACATGATGAGATATGTTCCACCTCAATTAAGTTGAGAGATGC 208
69 ahISleuAlaGlyInSerIleLeuSerGlyTyrrSerThrTyrrIleT 86
|||||
209 CCACATTAGTGGTCAAACATATGTTGCTGTCATCTACTATATATAT 258
86 yrrAlaIleAlaThrAlaProAsnMetPheAsnValAsnAspValleuGly 102
|||||
259 ATGTTATAGCACTGCACCCACATGTTTAACCTTAATGATGATTAAGG 308
103 ValTyrrSerProHisProTyrrGluGlnGluValSerAlaLeuGlyIle 119
|||||
309 GCATTACAGTCTCATCCAGATGAACAAGATTTCTGCTTAGGGGAT 358
119 eProTyrrSerGlnIleTyrrGlyTyrrTyrrArgValAsnPheGlyValIle 136
|||||
359 TCCATCTCCCAATATATGATGATGATGATGATGATGATGATGATGATG 408
136 spGluArgLeuHisArgAsnArgGluTyrrArgAspArgTyrrTyrrArgAsn 152
|||||
409 ATGACAAATTAATCATGTAATAGGGCTACAGATAGATGATTAACAGTA 458
153 LeuAsnIleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPheProP 159
|||||
459 TTAGATATTCCTCCAGCAGCAGATGTTAGTGGCAGGTTCCCTCC 508
169 oAspHisGlnAlaTrpArgGluGluTyrrProTrpIleHisAlaProGln 186
|||||
509 GGAGCATAGAGCTTGGAGGGAAGACCGTGATTCATCATGACACGGCGG 558
186 lYcysGlyAsnSerSerArgThr 193
|||||
559 GTTGTGGGAATGCTCCAGATCA 581

```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74430

seq_documentation_block:

ID AAF74430 standard; DNA: 564 BP.
 XX
 AC AAF74430;

```

XX 08-MAY-2001 (first entry)
DT
XX
XX Vibrio cholerae ctxA gene 564 bp region nucleotide sequence.
DE
XX
XX Vibrio cholerae; ctxA gene; detection; identification; toxigenic; ds.
KW
XX
XX Vibrio cholerae.
OS
XX
XX ZA200001349-A.
XX
XX 27-DEC-2000.
XX
XX 16-MAR-2000; 2000ZA-0001349.
XX
XX 16-MAR-2000; 2000ZA-0001349.
XX
XX (UYPR-) UNIV PRETORIA.
XX
XX Theron J, Venter SN, Brozel VS, Du Preez M;
XX
XX WPI, 2001-191829/19.
XX
XX Novel oligonucleotide primer useful for PCR amplification of toxigenic
PT Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
PT organisms present in the sample
XX
XX Example 1; Fig 6; 44pp; English.
XX
XX The present invention describes a method for amplifying (A) toxigenic
XX Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
XX reaction (PCR), by combining a medium containing toxigenic V. cholerae
XX nucleic acid with an oligonucleotide primer (I) which will bind to a
XX section of V. cholerae nucleic acid, in an environment suitable for
XX carrying out PCR; (b) combining the medium with a second primer which
XX will bind downstream to the first primer to a section of the
XX complementary strand of the V. cholerae nucleic acid; and (c) amplifying
XX corresponding sections of V.cholerae nucleic acid present in the medium.
XX The method is useful for detecting toxigenic V. cholerae organisms or
XX toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
XX water sample e.g., tap water sample, river water sample, lake water
XX sample or a sewage sample. The method is useful in PCR assays for
XX detecting a microorganism or nucleic acid of the microorganism in a
XX sample. (1) and amplification products are useful for detecting toxigenic
XX V. cholerae nucleic acid in a sample. The PCR method involving (I) is
XX rapid, economical, specific and sensitive for detecting toxigenic
XX V. cholerae in water samples. The present sequence represents a
XX region from the V. cholerae ctxA gene, which is used in an example from
XX the present invention.
XX
XX Sequence 564 BP; 159 A; 105 C; 139 G; 161 T; 0 other;

```

alignment_scores:

Quality:	910.00	Length:	187
Ratio: <td>5.056 <td>Gaps: <td>0</td> </td></td>	5.056 <td>Gaps: <td>0</td> </td>	Gaps: <td>0</td>	0
Percent Similarity: <td>96.257 <td>Percent Identity: <td>87.166</td> </td></td>	96.257 <td>Percent Identity: <td>87.166</td> </td>	Percent Identity: <td>87.166</td>	87.166

alignment_block:

US-09-528-682-1 x AAF74430 ..

Align seg 1/1 to: AAF74430 from: 1 to: 564

```

7 ArgAlaAspSerArgProPoaSgLuIleLysArgSerGlyLysLeu 23
|||||
1 CGGCGAGATTCTAGACCTCGATGAATTAAGCAGTCAAGTGTCTTAT 50
23 tProArgGlyHisAnGluTyrrPheAspArgGlyThrcGlnMetAsnIle 40
|||||
51 GCCAAGAGCAGAGTAGTACTTTGACCGAGGTACTCAATGAATATCA 100
40 snLeuTyrrAspHisAlaArgGlyThrcGlnThrcGlyPheValArgTyrrAsp 56
|||||

```



```

101 ACCTTATGATCATGCAAGAGAACTCAGACGGATTGTTAGGCACGAT 150
57 ASpglytyrValSerThrSerLeuSerAlaHisLeuAlaI 73
151 GATGGATATGTTCCACCTCAATTAAGTGAAGAGCCACCTAGGGG 200
73 yGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaI 90
201 TCAACATATATGTCGTCATCTCTACTTATATATATATATATATACCA 250
90 hAlaIProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerPro 106
251 CTGCACCCACATGTTTACGTTAATGATGATATATAGGCGCATACAGCTC 300
107 HisProTyrGlnGlnGluValSerAlaLeuGlyGlyIleProTyrSerG 123
301 CATTCACATGACACAGAAAGTTTCTGCTTTCAGTGGGATTCATACCTCCA 350
123 nIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuH 140
351 AATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
140 iSaTgAsnAlaGlyTyrArgAspArgTyrTyrArgAsnLeuAlaI 156
401 ATCTAATATAGGCGCTACAGATGATGATGATGATGATGATGATGATGCT 450
157 ProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAl 173
451 CCACGACAGAGTGGTATGATGATGATGATGATGATGATGATGATGATGAT 500
173 aTTPArgGluGluProTyrIleHisIleAlaProGlnGlyCysGlyAsnS 190
501 TTGGAGGGAAGAGCCGATGATGATGATGATGATGATGATGATGATGATG 550
190 eTserArgThr 193
551 CTCGACGATCA 561

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74431
seq_documentation_block:
ID AAF74431 standard; DNA; 347 BP.
XX
AC AAF74431;
XX
DT 08-MAY-2001 (first entry)
XX
DE Vibrio cholerae ctxA gene 347 bp region nucleotide sequence.
XX
KM Vibrio cholerae; ctxA gene; detection; identification; toxigenic; ds.
XX
OS Vibrio cholerae.
XX
PN ZA200001349-A.
XX
PD 27-DEC-2000.
XX
PF 16-MAR-2000; 2000ZA-0001349.
XX
PR 16-MAR-2000; 2000ZA-0001349.
XX
PA (UYPR-) UNIV PRETORIA.
XX
PI Theron J, Venter SN, Brozel VS, Du Preez M;
XX
DR WPI; 2001-191829/19.
XX
XX
PT Novel oligonucleotide primer useful for PCR amplification of toxigenic
PT Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
PT organisms present in the sample
PS Example 2; Fig 7; 4pp; English.
XX

```

```

CC The present invention describes a method for amplifying (A1) toxigenic
CC Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
CC reaction (PCR), by combining a medium containing toxigenic V. cholerae
CC nucleic acid with an oligonucleotide primer (1) which will bind to a
CC section of V. cholerae nucleic acid, in an environment suitable for
CC carrying out PCR; (b) combining the medium with a second primer which
CC will bind downstream to the first primer to a section of the
CC complementary strand of the V. cholerae nucleic acid; and (c) amplifying
CC corresponding sections of V. cholerae nucleic acid present in the medium.
CC The method is useful for detecting toxigenic V. cholerae organisms or
CC toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
CC water sample e.g., tap water sample, river water sample, lake water
CC sample or a sewage sample. The method is useful in PCR assays for
CC detecting a microorganism or nucleic acid of the microorganism in a
CC sample. (1) and amplification products are useful for detecting toxigenic
CC V. cholerae nucleic acid in a sample. The PCR method involving (1) is
CC rapid, economical, specific and sensitive for detecting toxigenic
CC V. cholerae in water samples. The present sequence represents a
CC region from the V. cholerae ctxA gene, which is used in an example from
CC the present invention.
SQ
Sequence 347 BP; 101 A; 66 C; 77 G; 103 T; 0 other;

alignment_scores:
Quality: 562.00 Length: 115
Ratio: 4.930 Gaps: 0
Percent Similarity: 99.130 Percent Identity: 91.304

alignment_block:
US-09-528-682-1 x AAF74431 ..
Align seg 1/1 to: AAF74431 from: 1 to: 347

7 ArgAlaAspSerArgProProAspGluIleLysArgSerGlyGlyLeu 23
1 CGGGCAGATTCTAGACCTCCTGATGAATAAAGCAGTCAGTGTCTTAT 50
23 tProAlaGlnHisAsnGluTyrPheAspArgGlyTyrGlnMetAsnIleA 40
51 GCCAAGAGCAGAGTGAAGTCTTTCAGCGAGTACTCAATGCAATATCA 100
40 snLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAsp 56
101 ACCTTATGATCATGCAAGAGAACTCAGACGGATTTGTTAGCAGAT 150
57 ASpglytyrValSerThrSerLeuSerAlaHisLeuAlaI 73
151 GATGGATATGTTCCACCTCAATTAAGTGAAGAGCCACCTAGGGG 200
73 yGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaI 90
201 TCAACATATATGTCGTCATCTCTACTTATATATATATATATATACCA 250
90 hAlaIProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerPro 106
251 CTGCACCCACATGTTTACGTTAATGATGATATATAGGCGCATACAGCTC 300
107 HisProTyrGlnGlnGluValSerAlaLeuGlyGlyIleProTyr 121
301 CATTCACATGACACAGAAAGTTTCTGCTTTCAGTGGGATTCATACCTCCA 350

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAT03448
seq_documentation_block:
ID AAT03448 standard; DNA; 777 BP.
XX
AC AAT03448;
XX
DT 02-MAY-1996 (first entry)
XX
DE Shigella flexneri thermolabile enterotoxin gene fragment Lf.
XX

```


KW Thermolabile enterotoxin; subunit-A; subunit-B; detection; probe;
KM LfE fragment; enterobacteria; ss.
XX Shigella flexneri.
XX R02031948-C1.
XX PD 27-MAR-1995.
XX PF 11-DEC-1991; 91RU-0016860.
XX PR 11-DEC-1991; 91SU-4016860.
XX PA (NIZH-) NIZHEGOROD EPIDEMIOLOGY MICROBIOLOG INST.
XX PI Masepa VN, Skoblo LE, Ulanova TI;
XX DR WPI; 1995-343066/44.
XX PF LfE DNA fragment for recognition of thermolabile enterobacterium
PT enterotoxin gene - can be used as a probe for recognition of
PT Lf-toxin DNA gene
XX PS Claim 1; Columns 3-4; 5pp; Russian.
XX CC The present sequence is that of the LfE fragment which encodes part
CC of the thermolabile enterotoxin subunit-A sequence and all of the
CC subunit-B sequence of Shigella flexneri. The DNA fragment is useful
CC as a probe for specific detection of toxigenic enterobacteria.
XX SQ Sequence 777 BP; 288 A; 141 C; 153 G; 195 T; 0 other;

alignment_scores:
Quality: 354.00 Length: 68
Ratio: 5.206 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-528-682-1 x AAT03448 ..

Align seg 1/1 to: AAT03448 from: 1 to: 777

173 AATTPARGUGLUPROTTPLEHSHSALAPROGLNGLYCYSGLYAS 189
|||||
3 GCTTGAGAGAAACCTGATTCATGACCAAGGTTGGA 52
189 nserSArGTThrIleThrGlyAspThrcYsAnGLuGluThrGlnAsnL 206
|||||
53 TTCATCAAGACATTCACAGGTGATCTGTATGAGAGGCCAGAAATC 102
206 euserThrIleTyrLeuArgGluTyrGlnSerLysValLysArgGlnIle 222
|||||
103 TAGACACATATATCTCAGGGAATATCAATCAAAAGTTAAAGAGCAATA 152
223 PhesArSPTyrGlnSerGluValAspIleTyrAsnArgIleArgSpL 239
|||||
153 TTTTCGACTATCATCGACAGGTTGACATATATAACAGAAATCGGATGA 202

239 uleu 240
||||
203 ATTA 206

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62460

seq_documentation_block:

ID AAV62460 standard; DNA: 2385 BP.

XX AAV62460;

XX 15-FEB-1999 (first entry)

DE Helicobacter pylori ureb-Vibrio cholerae A2+B subunit gene fusion.

XX Vaccine; antigen: Ureb; antigen; toxin; diagnosis; gastritis;
KM ulcer; stomach cancer; ss.
XX OS Chimeric - Helicobacter pylori.
XX OS Chimeric - Vibrio cholerae.
XX Key Location/Qualifiers
FT CDS 1..1680
FT /*tag= a
FT /product= ureb
FT CDS 1681..2256
FT /*tag= b
FT /product= A2+B subunits

XX WO9844130-A1.
XX 08-OCT-1998.
XX 31-MAR-1998; 98WO-KR00073.
XX 31-MAR-1997; 97KR-0011951.
XX 31-MAR-1997; 97KR-0011950.
XX (DAEW-) DAEWOONG PHARM CO LTD.
XX Chol D, Jung H, Kim B, Park M, Shin S, Yu Y;
XX WPI; 1998-568279/48.
XX P-PSDB: AAW80599.

PT New chimeric proteins for use against Helicobacter pylori -
PT comprising an antigenic protein of H. pylori and A1 and B subunits
PT of Vibrio cholerae toxin, preferably produced by recombinant
PT techniques

PS Claim 3; Page 77-78; 102pp; English.

XX This is the nucleotide sequence of a fusion gene prepared by
CC ligating the ureb gene of Helicobacter pylori and the A2 and B
CC subunits genes of Vibrio cholerae toxin. It codes for a claimed
CC 749-amino acid fusion protein (see AAW80599). The invention relates
CC to: chimeric proteins comprising antigenic proteins of H. pylori
CC and A2 and B subunits of V. cholerae toxin; recombinant DNAs
CC encoding such chimeric proteins; recombinant expression vectors;
CC a process for preparing the chimeric proteins using transformed
CC microbial host cells; and to preventative and therapeutic vaccines
CC comprising the chimeric proteins for H. pylori-associated diseases
CC such as gastritis, gastric ulcer, duodenal ulcer and gastric cancer
CC (claimed). The chimeric proteins are designed to have excellent
CC immunogenicity, to be stable in the stomach, to penetrate the
CC mucous membrane of the intestine, and to stimulate production of
CC sIgA. They can additionally be used as active ingredients in
CC diagnostic kits for H. pylori infection, and for production of
XX anti-H. pylori antibody.

SQ Sequence 2385 BP; 779 A; 486 C; 509 G; 611 T; 0 other;

alignment_scores:
Quality: 252.00 Length: 112
Ratio: 3.111 Gaps: 3
Percent Similarity: 72.321 Percent Identity: 47.321

alignment_block:
US-09-528-682-1 x AAV62460 ..

Align seg 1/1 to: AAV62460 from: 1 to: 2385

137 GUARGLeuHISARGsAnArgGluTyrArgAspArgTyrTyrArg..... 151
::::: ||| ::::|
1557 CAGAAATATCATCAAAAAGACATGCAATTCACAGACATCAACGCGTCACA 1606


```
152 .....AsnLeuAsnIleAlaProAlaGluAspGlyTyrA 163
      |||||
1607 TTGAAGTCAATTCTGAACCTA.....CCATGCTGTCGTGATGGC 1647
      |||||
163 rGleuAlaGlyPheProAspHisGlnAla.TTPArgGluGluProTr 179
      ::::: ||| ::::: |||||
1648 AAAGAAGTAACTCTAAACCAATAAAGTGAGATTGGAAGAGCCGTG 1697
      ::::: ||| ::::: |||||
179 PLeuHisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrG 196
      ||||| ||||| ||||| ||||| ::::: |||||
1698 GATTCATCATGACCGCGCGGTGTGGGAATGCTCCAGATCATGATCA 1747
      ::::: ||||| ::::: |||||
196 lAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArg 212
      ::::: ||||| ::::: ||||| ::::: |||||
1748 GTAATACTTGCATGAAAAAACCAAGCTAGGTGTAAATTCCTTGAC 1797
      ::::: ||||| ::::: |||||
213 GluTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGlnSerG 229
      ||||| ||||| ||||| ||||| |||||
1798 GAATACCAATCTAAAGTTAAAGACAAATATTTCAGGCTATCAATCTGA 1847
      ::::: ||||| ::::: |||||
229 uValAspIleTyrAsnArgIleArgAspGluLeu 240
      ::::: ||||| ::::: |||||
1848 TATTGATACACATTAATAGATTAGATGATTA 1881
```


gb_est1:AI102743	-	79.00	143.45	96.21	594	AI102743	EST212032	Normalized	+	77.00	139.17	166.42	604	BF575414	602135645F1	NIH_MGC
gb_est1:AM210136	+	79.00	142.25	98.62	604	AM210136	u151c10.y1	Rashpass	+	77.00	139.08	168.47	609	BI605457	RH71079.5	Sprtime RH D
gb_est2:BM285021	+	79.00	142.84	103.99	626	BM285021	K930910.y1	Ascaris	+	77.00	139.06	168.88	610	BI605330	RH63081.5	Sprtime RH D
gb_est1:BB574178	+	79.00	141.93	116.80	719	BB574178	BB521738	Riken full-16	+	77.00	138.76	175.49	626	AI125949	LP02919.5	Sprtime LP D
gb_est2:BG574777	+	79.00	141.24	127.70	719	BG574777	EST496155	GSTS	+	77.00	138.32	185.56	650	BI600859	BOGC719F	BOGC Bras
gb_est2:BG35038	+	79.00	140.91	132.47	737	BG35038	602457485F1	NIH_MGC_16	+	77.00	137.97	194.09	670	AG014025	Homo sapiens	genoml
gb_est1:AI684897	+	79.00	140.19	146.02	787	AI684897	wa86h08.x1	Soares	+	77.00	137.96	194.52	671	AM652302	EST320109	tomto ra
gb_gss:AZ934810	+	79.00	140.16	146.57	789	AI2934810	BT_Ba000202	Mr. Japc	+	77.00	137.22	213.73	715	AM652302	Sheared DNA-6k2	TR D
gb_est2:BI227110	+	79.00	139.93	151.01	805	BI227110	602952296F1	NIH_MGC_6	+	77.00	136.35	239.02	771	BF630594	HVMBB0012A06F	Hord
gb_est2:BE780666	+	79.00	139.76	257.91	1155	BE780666	601469344F1	NIH_MGC_6	+	77.00	135.94	252.00	799	BG418932	6025599912F1	NIH_MGC
gb_est2:BI685336	+	79.00	134.42	305.95	1296	BI685336	603382377F1	NCI_CGAP	+	77.00	134.77	292.76	884	BF280560	601811768F1	NIH_MGC
gb_est1:AI199448	+	78.50	147.59	56.54	357	AI129448	zn85h03.s1	Stratagene	+	77.00	134.01	322.70	944	BE780560	601582575F1	NIH_MGC
gb_est2:BG959506	+	78.50	144.49	84.18	497	BG959506	MR4-FH1052-150201	-001	+	77.00	133.23	356.72	1010	BE616292	6028443507F2	NIH_MGC
gb_est1:AI555861	+	78.50	141.26	127.33	657	AI528996	u43e02.y1	Soares	+	77.00	132.94	369.89	1035	AC063739	Pan troglodytes	10 da
gb_est2:BI17246	+	78.50	140.33	143.45	712	BI147246	BI147246	unpublished	+	77.00	131.84	426.32	1139	AC063739	Pan troglodytes	DN
gb_est2:BI155441	+	78.50	140.07	148.26	728	BI155441	BI155441	unpublished	+	77.00	127.12	780.85	1713	BI552261	603195922F1	NIH_MG
gb_est2:BG587737	+	78.50	139.98	150.08	734	BG587737	EST489512	MHAM	+	76.50	134.42	306.07	834	BC701267	602682939F1	NIH_MGC
gb_gss:BM533375	+	78.50	139.68	155.87	753	BM533375	BOHB2307F	BOHB	+	76.50	137.85	197.18	620	AI2653478	IM052707F	Mouse 10
gb_est2:BG914788	+	78.50	138.38	184.28	843	BG914788	602813589F1	NIH_MGC_16	+	76.50	136.10	246.64	721	BG159770	de86d11.y1	Kirsche
gb_est2:BG440651	+	78.50	137.97	194.09	875	BG490651	602520125F1	NIH_MGC_18	+	76.50	135.12	279.79	785	BM045368	603623364F1	NIH_MGC
gb_est2:BG414440	+	78.50	137.06	218.30	945	BI181562	UNL-P-FN-a0-f-09-0-UNI		+	76.50	134.42	306.07	834	BC701267	602682939F1	NIH_MGC
gb_est2:BI181562	+	78.50	136.83	224.84	964	BE299664	600944327F1	NIH_MGC_19	+	76.50	133.93	325.87	870	AL537110	AL537110	LIT_F1013
gb_est2:BE299664	+	78.50	135.19	227.50	1111	BE299664	602363245F1	NIH_MGC_19	+	76.50	132.81	376.50	959	BG532489	602366244F1	NIH_MGC
gb_est2:BF230828	+	78.50	127.88	708.33	2090	BF230828	601867526F1	NIH_MGC_1	+	76.50	130.77	489.07	1144	BC975930	602843041F1	NIH_MG
gb_est1:BB840688	+	78.00	147.16	59.72	361	BB840688	BB840688	Riken full-16	+	76.50	128.52	652.85	1390	AC097062	Pan troglodytes	DN
gb_gss:AO909105	+	78.00	145.86	70.56	548	AO909105	GSTC07539	Trypansoma	+	76.50	127.03	790.20	1581	BC015622	Homo sapiens	Stnl
gb_est1:AA803232	+	78.00	142.34	110.90	548	AA803232	GM0436.5	Sprtime GM Dros	+	76.50	127.03	790.20	1581	BC015622	Homo sapiens	Stnl
gb_est1:AI1486065	+	78.00	141.86	117.87	571	AI1486065	EST244386	tomato ovary	+	76.50	118.68	2.3e+03	3255	AK014599	Mus musculus	0 day
gb_est2:BI306660	+	78.00	141.82	118.48	573	BI306660	NL_5.124	Drought stres	+	76.50	118.32	2.4e+03	3357	AK005011	Mus musculus	adult
gb_est1:AA802683	+	78.00	141.60	121.87	584	AA802683	GM06110.5	Sprtime GM Dros	+	76.50	139.67	156.14	485	AA682866	J16b04.y1	Soares
gb_est2:BG655776	+	78.00	141.37	125.60	596	BG653776	SD13264.5	Sprtime SD Dros	+	76.00	138.22	187.16	548	BM277429	Tm-ad.A0A06.SRPL	TR R
gb_est2:BI658554	+	78.00	141.35	125.91	597	BI638554	SD20866.5	Sprtime SD Dros	+	76.00	138.22	188.14	550	AO964439	RPCT-23-322P16	TJ R
gb_est2:BI206448	+	78.00	141.33	126.23	598	BI206448	EST524488	ctfcs	+	76.00	137.99	193.76	561	AO968780	RPCT-23-322P16	TJ R
gb_est2:BG640828	+	78.00	141.31	126.34	599	BG640828	SD11981.5	Sprtime SD Dros	+	76.00	137.62	203.05	579	AO383867	RPCT11-138112	TV RP
gb_est2:BI113506	+	78.00	141.19	128.42	605	BI173506	RE16438.5	Sprtime RE Dros	+	76.00	137.21	214.07	600	AW597203	EST318826	MHAM
gb_est2:BI151326	+	78.00	141.14	129.37	608	BI163236	RE02704.5	Sprtime RE Dros	+	76.00	136.48	235.03	639	AG139018	Pan troglodytes	full
gb_est2:BI581255	+	78.00	141.12	139.69	609	BI581255	RH18623.5	Sprtime RH Dros	+	76.00	136.37	238.31	645	BE792005	601583240F1	NIH_MGC
gb_est1:AI1113376	+	78.00	140.87	133.81	622	AI113376	GH08445.5	Sprtime GH Dros	+	76.00	136.12	246.02	659	BE792005	601583240F1	NIH_MGC
gb_est2:BM318921	+	78.00	139.96	150.40	673	BM318921	K162806.y1	Ascaris	+	76.00	135.99	250.46	667	BG847824	1024019604.y1	C. re
gb_est1:AI1234208	+	78.00	139.89	151.73	677	AI294208	LP07529.5	Sprtime LP Dros	+	76.00	135.97	251.01	668	BB648070	BB648070	Riken full
gb_est1:AM037263	+	78.00	139.88	152.06	678	AM037263	EST274113	tomato mixed	+	76.00	135.69	259.98	684	BJ041966	BJ041966	NIBH
gb_est2:BF479279	+	78.00	139.64	156.74	692	BF479279	I48-294275	Ice plant	+	76.00	135.59	263.37	690	BB648001	BB648001	Riken full
gb_est1:AL653906	+	78.00	139.41	161.46	706	AL653906	AL653906	XGC	+	76.00	135.54	265.07	693	BI543385	SLA.C6	Sugar
gb_est2:BF308520	+	78.00	138.85	173.48	741	BF308520	601887667F1	NIH_MGC_17	+	76.00	135.49	266.77	696	AI598084	TS06d10.y1	NCI_CGAP
gb_est2:BG972287	+	78.00	138.21	188.25	783	BG972287	602841418F1	NCI_CGAP	+	76.00	135.26	274.77	710	BM300036	MCR054D08	24600 ice
gb_gss:AG847944	+	78.00	137.51	205.99	832	BG847944	1024019H01.x1	C. rein	+	76.00	135.23	275.92	712	AM12834	JALSL3C16	SHC
gb_gss:AG187693	+	78.00	131.65	436.71	1381	AG187693	Pan troglodytes	DNA	+	76.00	135.07	281.69	722	AM12834	JALSL3C16	SHC
gb_est2:BE411552	+	77.50	143.78	92.20	443	BE411552	ISC006.G09F90629	ITBC	+	76.00	134.91	287.49	732	BF458735	tm24601.y1	Soares
gb_gss:AO653882	+	77.50	139.95	150.70	617	AO653882	Sheared DNA-9N16	TF SH	+	76.00	134.74	293.92	743	AI478735	tm24601.y1	Soares
gb_est2:BI1833876	+	77.50	139.74	155.93	630	AI1833876	605096F05.x2	GF 1	+	76.00	134.57	300.40	754	BM048735	603625338F1	NIH_MGC
gb_gss:BM398288	+	77.50	136.74	175.97	665	BM398288	AG-ND-138P23	TF 1	+	76.00	134.51	302.66	758	BM044656	603625338F1	NIH_MGC
gb_est2:BF107273	+	77.50	138.55	160.18	666	BF107273	601824167F1	NIH_MGC_73	+	76.00	133.47	345.75	829	BI250850	602993320F1	NCI_CGA
gb_est1:AI200892	+	77.50	137.64	196.14	737	BI375192	601228346F1	NIH_MGC_16	+	76.00	133.75	350.09	832	BI250850	602993320F1	NCI_CGA
gb_est2:BE375192	+	77.50	137.64	196.14	737	BE375192	601228346F1	NIH_MGC_16	+	76.00	133.75	350.09	832	BI250850	602993320F1	NCI_CGA
gb_est1:AI200892	+	77.50	137.64	196.14	737	BI375192	601228346F1	NIH_MGC_16	+	76.00	133.75	350.09	832	BI250850	602993320F1	NCI_CGA
gb_est2:BI6564800	+	77.50	136.08	247.44	862	BI6564800	603288215F1	NCI_CGAP	+	76.00	131.92	421.85	948	BF025712	601670597F1	NIH_MGC
gb_gss:CMS01M17	+	77.50	135.66	261.18	894	AI150348	Anopheles	gambie	+	76.00	130.50	506.20	1072	BM584844	AGENCORFT	641078
gb_est2:BG9860735	+	77.50	132.39	397.16	1186	BG9860735	6028836610F1	NCI_CGAP	+	76.00	132.24	1.5e+03	2189	AK003586	Mus musculus	18 da
gb_gss:AO910494	+	77.00	143.57	94.71	413	AO910494	GSTC05730	Trypansoma	+	75.50	139.52	159.25	450	AZ097806	RPCT-23-475A22	TV R
gb_est1:AA033034	+	77.00	142.10	114.36	469	AA033034	m137805.y1	Soares	+	75.50	139.16	189.51	506	BI095243	dac03a29.y1	NICH
gb_est2:BE455520	+	77.00	141.97	116.18	474	BE455520	WHE1287-1290	BOB25	+	75.50	137.78	199.03	523	AI7020240	LJNEST1E7F	Lotus
gb_est2:BG711191	+	77.00	141.54	122.78	492	BG711191	602719609F1	NIH_MGC_60	+	75.50	137.62	202.99	530	BE011850	130542	MARC
gb_gss:TM2960030	+	77.00	140.84	134.42	523	AI485322	T. brucei	sheared	+	75.50	136.99	220.26	560	BM094774	sa121b07.y1	Gm-cl06
gb_est2:BI333851	+	77.00	139.92	151.13	566	BI333851	GM25375	Sprtime GM Dros	+	75.50	136.99	220.26	560	BM094774	sa121b07.y1	Gm-cl06
gb_est1:AI297329	+	77.00	139.52	159.12	566	AI297329	LP11607	Sprtime LP Dros	+	75.50	136.05	248.22	607	BE992042	UT-M-B21	bee-d-03-0
gb_est1:AI257564	+	77.00	139.46	160.33	569	AI257564	LP05968	Sprtime LP Dros	+	75.50	136.02	249.44	609	BE992042	UT-M-B21	bee-d-03-0
gb_est1:AI1238573	+	77.00	139.44	160.73	590	AI257564	GH14642	Sprtime GH Dros	+	75.50	136.02	249.44	614	BI066942	BI066942	NIBH
gb_est2:BG640447	+	77.00	139.42	161.14	591	BI6640447	SD11465.5	Sprtime SD Dros	+	75.50	134.49	307.67	632	BF487338	ATY22185	Sprtime AT D
gb_est2:BG656546</																

gb_est2:BF495435	75.50	132.52	390.56	824	BF495435	AT04362.5ptime	AT Dros	gb_est2:BG104639	75.00	130.74	490.73	880	BG104639	602311329F1	NIH_MGC	
gb_est2:BG495624	75.50	130.74	490.61	961	BG495624	602539959F1	NIH_MGC	gb_est2:BG579934	75.00	129.89	547.14	947	BF579934	602095514F1	NIH_MGC	
gb_gss:CMS0502P	-	75.50	524.29	1005	AL214882	Tetradon nigr	NIH_MGC	gb_est2:BG038475	75.00	129.83	587.89	994	BG038475	602325508F1	NIH_MGC	
gb_est2:BE1616366	75.50	129.92	575.35	1070	BE161636	601278172F1	NIH_MGC	gb_est2:BE121150	75.00	129.31	589.65	996	BE121150	601650791F1	NIH_MGC	
gb_est2:BF559968	75.50	127.92	704.87	1227	BF559968	602186135F1	NIH_MGC	gb_est2:BE177453	75.00	129.12	597.56	1005	BE177453	601314068F1	NIH_MGC	
gb_est2:BF979323	75.50	126.88	805.02	1342	BF979323	602241502F1	NIH_MGC	gb_est2:BE174584	75.00	129.16	601.09	1009	BE174584	602314068F1	NIH_MGC	
gb_est2:BF038306	75.50	124.60	805.02	1342	BF038306	601458125F1	NIH_MGC	gb_gss:AC073926	75.00	127.91	705.43	1124	AG073926	601570894F1	NIH_MGC	
gb_est1:BE1217683	75.00	141.70	120.30	165	BE121768	DEP1432	Rat Lambda	gb_est2:BE130763	75.00	126.44	852.41	1277	BE130763	601570894F1	NIH_MGC	
gb_est2:BG4100295	-	75.00	139.63	156.97	408	BG410095	947051A01	xi 947 - 2	gb_est2:BA225509	75.00	120.73	186.03	2091	BM425509	AGNCNCOURT	6139439
gb_est2:BI952721	75.00	139.29	163.86	420	BI952721	HVSEMO07116F	Hordeum	gb_est1:AV224661	74.50	143.13	100.20	276	AV224661	AV224661	Riken full	
gb_est1:AA559232	75.00	138.62	178.53	445	AA559232	nm1A07	sl NCI_CGAP	gb_est1:BB808141	74.50	139.90	151.67	365	BB808141	BB808141	Riken full	
gb_est2:BF077267	75.00	138.50	181.51	450	BF077267	2271104	MARC	2BOV Bos t	74.50	137.89	195.07	463	AA076589	zm91912	xi Strategic	
gb_est2:BG898378	75.00	138.07	188.74	462	BF077267	UI-M-BH3-ue-c	-02-0-UI	gb_est2:BA969473	74.50	137.15	212.80	464	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt	74.50	132.74	379.93	678	BE946839	602080732F2	NIH_MGC	
gb_est2:BF849109	75.00	136.63	230.71	529	BI801116	H040G07	Endosperm	lbdt								

gb_est1:AM980082	74.00	133.48	345.30	582	AM980082	EST341589	tomato	root	gb_est2:BF183697	73.50	120.78	1.8e+03	1599	BF183697	6018069595R1	NH_MGC		
gb_est1:AT142102	74.00	133.26	355.91	593	AT1542102	SD08184	Sprtime	SD	gb_est2:BF133349	73.50	118.87	2.2e+03	1885	BF133349	601645337R1	NH_MGC		
gb_est2:BM141227	74.00	133.25	355.92	594	BM014227	603639930F1	NH_MGC	87	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_est2:BF133034	74.00	133.11	362.15	601	BF163034	RM594775	Sprtime	RH	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_est2:BF1171064	74.00	133.03	362.15	605	BF1171064	REL12735	Sprtime	RE	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_est1:AA146554	74.00	132.84	374.73	615	AA146554	z043604	x1	Stratagene	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_est1:AM070653	74.00	132.69	381.98	623	AM070653	MR3-RT0487	-990100	-103	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_gss:AG068493	74.00	132.55	389.27	631	AG068493	RCPC193-DpnII	-2505	TV	gb_est2:BF133349	73.00	126.68	229.14	370	BF133349	601645337R1	NH_MGC		
gb_gss:AG098558	74.00	131.66	435.88	681	AG098558	Pan	trogloclites	DNA	gb_est1:AM330403	73.00	132.86	374.16	515	AM330403	EST340776	tomato	fr	
gb_gss:AG128823	74.00	131.60	439.68	685	AG128823	Pan	trogloclites	DNA	gb_est1:AM330403	73.00	132.86	374.16	515	AM330403	EST340776	tomato	fr	
gb_gss:AG128237	74.00	131.53	443.49	689	AG128237	qk2209	x1	NCL	gb_est1:AM330403	73.00	132.86	374.16	515	AM330403	EST340776	tomato	fr	
gb_gss:AG2019847	74.00	131.53	443.49	689	AG2019847	RCPC1-23	-29915	TV	gb_gss:AG2019847	73.00	132.86	374.16	515	AG2019847	RCPC1-23	-29915	TV	
gb_est1:AM318852	74.00	131.35	454.03	700	AM318852	un08412	x1	Sugeno	gb_est2:BM127115	73.00	132.05	414.70	552	BM127115	GH13728	Sprtime	GH	
gb_est2:BF1250056	74.00	129.67	592.82	838	BF1250056	602995406F1	NH_MGC	87	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_gss:BM074321	74.00	129.09	606.58	851	BM074321	RCPC1-24	-303420	TV	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BF109321	74.00	129.05	609.76	854	BF109321	60289711F1	NCL	CGAP	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BF573762	74.00	128.79	629.98	873	BF573762	60207073F2	NH_MGC	87	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_gss:CN50780K	74.00	128.44	659.09	900	AL433854	T3	end	of	clone	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87
gb_gss:BM11658	74.00	128.39	663.44	904	BM11658	ENTTS587F	Entamoeba	h1	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BG031541	74.00	127.86	709.66	946	BG031541	60229337F2	NH_MGC	87	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BF573177	74.00	127.86	709.66	946	BF573177	601652082R1	NH_MGC	87	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BM799531	74.00	125.92	910.36	1119	BM799531	AGENCOURT	6464948	NH	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est2:BF563343	74.00	120.68	1.8e+03	1761	BF563343	601335284F1	NH_MGC	3	gb_est2:BF1250056	73.00	132.05	414.70	552	BF1250056	602995406F1	NH_MGC	87	
gb_est1:BM054439	73.50	142.19	113.07	251	BM054439	BM054439	Riken	full	gb_est1:BM054439	73.00	130.09	533.26	653	BM054439	BM054439	Riken	full	
gb_est1:BM187073	73.50	138.61	178.90	342	BM187073	rs20H01	x1	Sommer	gb_gss:BM187073	73.00	129.92	545.40	664	BM187073	rs20H01	x1	Sommer	
gb_est2:BG5977038	73.50	137.48	206.70	377	BG5977038	HOA33-1	-110	HOA	gb_est2:BG5977038	73.00	129.92	545.40	664	BG5977038	HOA33-1	-110	HOA	
gb_est2:BE517149	73.50	137.03	219.02	392	BE517149	WHE0624	G05	N102A	gb_est2:BE517149	73.00	129.92	545.40	664	BE517149	WHE0624	G05	N102A	
gb_est1:AM945059	73.50	135.77	257.32	437	AM945059	UT-H-B12	-ahf	-g-08	-0-UT	gb_est2:BE517149	73.00	129.92	545.40	664	BE517149	WHE0624	G05	N102A
gb_est2:BF1442984	73.50	135.77	257.32	437	BF1442984	ds192607	x1	NICHD	gb_est2:BE517149	73.00	129.92	545.40	664	BE517149	WHE0624	G05	N102A	
gb_est1:BM747572	73.50	135.64	261.69	442	BM747572	BM747572	Riken	full	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM200485	73.50	134.86	289.37	473	AM200485	ds22a08	x1	normalized	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM200304	73.50	133.94	325.44	512	AM200304	QV2-ST0296	-150200	-028	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AL1162617	73.50	133.74	333.96	521	AL1162617	A020P34U	Hybrid	aspen	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est2:BF078225	73.50	133.74	333.96	521	BF078225	228642	MARC	2P1G	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est2:BF078225	73.50	133.74	333.96	521	BF078225	228642	MARC	2P1G	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM064928	73.50	133.01	366.79	551	AM064928	un52601	T1	Soares	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM20233	73.50	133.87	373.67	562	AM20233	QV2-ST0296	-150200	-040	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est2:BG021521	73.50	133.72	380.59	569	BG021521	ds66e10	x1	Wellcome	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est2:BF160588	73.50	132.32	400.60	589	BF160588	387822	MARC	2P1G	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AA008559	73.50	132.05	414.80	604	AA008559	mg86c11	x1	Soares	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM607177	73.50	132.05	414.80	604	AM607177	mg86c11	x1	Soares	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:AM764425	73.50	131.82	427.10	615	AM764425	ds92801	x1	NCL	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_est1:BM35825	73.50	131.80	428.13	616	BM35825	GH13719	Sprtime	GH	gb_est2:BF133349	73.00	129.92	545.40	664	BF133349	REL12735	Sprtime	RE	
gb_gss:BM452057	73.50	131.20	462.58	619	BM452057	IM0251P10R	Mouse	10Kb	gb_gss:BM452057	73.00	126.82	811.44	867	BM452057	IM0251P10R	Mouse	10Kb	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	
gb_est2:BM299704	73.50	131.06	471.06	657	BM299704	L48-191973	Ice	plant	gb_est2:BM299704	73.00	126.82	811.44	867	BM299704	L48-191973	Ice	plant	

gb_hlc:AK009776	-	73.00	119.13	2.2e+03	1688	AK009776	Mus musculus adult m	gb_est2:BF248206	+	72.50	126.23	874.73	836	BF248206	601821205F1 NIH_MGC
gb_hlc:AK009794	-	73.00	119.13	2.2e+03	1688	AK009794	Mus musculus adult m	gb_gss:BF679143	+	72.50	126.19	879.39	839	AZ679143	ENTL78TR Entomoba
gb_hlc:AK009795	-	73.00	119.10	2.2e+03	1683	AK010215	Mus musculus adult m	gb_est2:BF507733	+	72.50	126.12	887.18	844	BF507733	60187744F1 NIH_MGC
gb_hlc:AK009346	-	73.00	119.07	2.2e+03	1686	AK009346	Mus musculus adult m	gb_gss:AZ135376	+	72.50	126.12	887.18	844	AZ135376	OSUNDB0115B231 CGCI
gb_hlc:AK0010154	-	73.00	119.05	2.2e+03	1686	AK010154	Mus musculus adult m	gb_est2:BF191963	+	72.50	125.77	933.75	873	BF191963	603311605F1 NCI_CGA
gb_hlc:AK009871	-	73.00	119.05	2.2e+03	1699	AK009871	Mus musculus adult m	gb_est2:BF572128	+	72.50	125.72	934.73	874	BF572128	602128469F1 NIH_MGC
gb_hlc:AK009619	-	73.00	119.03	2.2e+03	1703	AK009619	Mus musculus adult m	gb_est2:BF154798	+	72.50	125.58	951.83	885	BF154798	60182901F1 NIH_MGC
gb_hlc:AK0010157	-	73.00	118.87	2.2e+03	1727	AK010157	Mus musculus adult m	gb_est2:BF315081	+	72.50	125.56	955.43	886	BF315081	601902517F1 NIH_MGC
gb_est1:AM702721	-	72.50	138.16	189.47	298	AM702721	TGEST2286107.y1 TGRH+	gb_est2:BF783465	+	72.50	124.99	1.0e+03	931	BF783465	60177190F1 NIH_MGC
gb_est1:AM66496	+	72.50	137.85	197.06	307	AM66496	wk47C01.x1 NCI_CGAP_P	gb_est2:BF147949	+	72.50	124.92	1.0e+03	937	BF147949	601901696F1 NIH_MGC
gb_est1:AM661727	+	72.50	135.44	268.52	377	AM661727	h180a07.x1 NCI_CGAP_P	gb_gss:CMS0699A	+	72.50	124.98	1.0e+03	940	AL166522	73 end of Clone AX0
gb_est1:AL601188	+	72.50	135.40	277.01	385	AL601188	DKFPJ3J0D640.r1 313	gb_est2:BF316832	+	72.50	124.44	1.0e+03	944	BF127232	601650879F1 NIH_MGC
gb_est1:BE028644	-	72.50	134.61	299.62	405	BE029644	kp3a03.y1 TBN95TM-SSR	gb_est1:BF216710	+	72.50	124.72	1.1e+03	953	BE216710	HV_CEB0014J1F1 Hordeum
gb_est1:AM488045	-	72.50	134.56	300.81	405	AM488045	58930 MARC 4BOV Bos te	gb_est2:BF315285	+	72.50	124.61	1.1e+03	956	BF315285	601899159F1 NIH_MGC
gb_est2:BF713628	+	72.50	134.53	301.91	408	BF15628	mab09911.y1 Soares NMR	gb_gss:CMS064LA	+	72.50	124.61	1.1e+03	962	AL149876	T7 end of Clone AX0
gb_est2:RG607104	+	72.50	134.00	332.99	427	RG607104	WHE2489_D08_H15Z5 Trif	gb_est2:BF313777	+	72.50	124.56	1.1e+03	966	BF313777	601901126F1 NIH_MGC
gb_est2:RS1118	+	72.50	133.60	339.95	442	RS1118	y71a01.r1 Soares Infant	gb_est2:BF316832	+	72.50	124.44	1.1e+03	977	BF166332	601903345F1 NIH_MGC
gb_est2:BE464122	-	72.50	133.02	366.51	465	BE464122	hy20h12.x1 NCI_CGAP_GC	gb_est2:BF895749	+	72.50	123.91	1.2e+03	1022	BF975749	602246332F1 NIH_MGC
gb_est1:AM081352	-	72.50	132.84	374.73	472	AM081352	xc41b11.x1 NCI_CGAP_GC	gb_est2:BF8958051	+	72.50	123.71	1.2e+03	1045	BE898051	601440362F1 NIH_MGC
gb_gss:AM081352	-	72.50	132.84	374.73	472	AM081352	xc41b11.x1 NCI_CGAP_GC	gb_est2:BF346158	+	72.50	123.33	1.3e+03	1075	BF346158	602071618F1 NCI_CG
gb_est2:BF628187	+	72.50	132.48	392.52	487	BF628187	HVSMED0003F222 Hordeum	gb_est2:BF674994	+	72.50	123.04	1.3e+03	1102	BF475756	ACBNC0UR_6480254
gb_gss:AQ90618	+	72.50	132.46	393.72	488	AQ90618	RPCT-11-270H14.TV RCT	gb_est2:BF674994	+	72.50	122.87	1.3e+03	1118	BF674994	60184803F1 NIH_MG
gb_est2:BM428969	+	72.50	132.27	403.33	496	BM428969	952026D12.y1 952 - BM5	gb_est2:BF131633	+	72.00	136.39	237.78	318	BM502477	BB502477 RIKEN full
gb_est1:AV925043	+	72.50	132.15	403.37	501	AV925043	AV925043 K. Sato unpuh	gb_est1:AM461923	+	72.00	133.65	337.86	403	AM461923	BP230007A10H9 Soare
gb_est1:BF234115	+	72.50	131.97	419.10	509	BF234115	EST72137 Schistosoma	gb_est2:BA273434	+	72.00	133.57	341.60	406	AA273434	WHE2901.HO9_O17ZS W
gb_est2:BE85487	+	72.50	131.93	421.55	511	BE85487	7937b01.x1 NCI_CGAP_S	gb_est2:BA273434	+	72.00	133.51	344.10	408	BF803342	WHE2901.HO9_O17ZS W
gb_gss:TMV91A10P	+	72.50	131.46	447.49	532	AL459572	T. brucei sheared gend	gb_est2:BM410149	+	72.00	131.68	435.17	478	BM410149	EST584476 tomato br
gb_est1:AM153621	-	72.50	131.44	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.44	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621	-	72.50	131.46	448.74	533	AM149142	x1f10h09.x1 NCI_CGAP_K1	gb_est2:BF157379	+	72.00	130.86	483.25	513	BF157379	RM09288.5prtime RH D
gb_est1:AM153621															

gb_est2:BG034236	-	72.00	123.70	1.2e+03	953	1	BG034236	602302494Fl NIH_MGC_87	gb_est2:BM442159	-	71.50	126.05	895.58	712	1	BM442159	EBan01.SQ002.G03_R
gb_est2:BE260590	-	72.00	123.44	1.3e+03	905	1	BE260590	601146069Fl NIH_MGC_19	gb_est1:AE555533	-	71.50	125.92	910.54	720	1	AE555533	XGC_gastru
gb_est1:AL560901	+	72.00	123.11	1.3e+03	1003	1	AL560901	AL560901 LTI.NFL010E	gb_gss:BM104330	-	71.50	125.87	921.82	726	1	BM104330	BOCEP47F BOCE Bras
gb_est2:BF570467	+	72.00	123.02	1.3e+03	1011	1	BF570467	602186268Fl NIH_MGC_4	gb_est2:BI934986	-	71.50	125.67	940.71	736	1	BI934986	EST35475 Tomato Fl
gb_gss:ACNS0453	+	72.00	122.46	1.4e+03	1061	1	AL290496	Tetradon nigroviridis	gb_est2:BI761041	-	71.50	125.57	952.10	742	1	BI761041	603048581Fl NIH_MGC
gb_gss:ACNS0415	+	72.00	121.66	1.4e+03	1137	1	AG182715	Pan troglodytes DNA,	gb_gss:BG162884	-	71.50	125.54	955.91	744	1	BG162884	6143211Fl y1 Wellcome
gb_est2:BM106426	+	72.00	121.41	1.6e+03	1162	1	BI409837	602961945Fl NCI CGAP	gb_gss:AE2506171	-	71.50	125.40	973.11	753	1	AE2506171	IM0347704F Mouse 10
gb_gss:CM064937	-	72.00	121.39	1.6e+03	1164	1	AL411384	T3 end of clone ANOVA	gb_gss:AE231385	+	71.50	125.31	984.63	759	1	AE231385	OSJB0108P19F CUGI
gb_gss:AC110400	+	72.00	121.17	1.7e+03	1166	1	AG114400	Pan troglodytes DNA,	gb_est2:BG732200	-	71.50	125.14	1.0e+03	770	1	BG732200	602663158Fl NIH_MGC
gb_est2:AC002543	+	72.00	119.15	2.2e+03	1412	1	BE250243	600943234Fl NIH_MGC_1	gb_est2:BF106460	-	71.50	125.06	1.0e+03	776	1	BF106460	60182316Fl NIH_MGC
gb_hlc:AC002793	+	72.00	118.18	2.5e+03	1536	1	AK002793	Mus musculus adult m	gb_est2:BF025801	-	71.50	124.95	1.0e+03	783	1	BF025801	601670222Fl NIH_MGC
gb_est2:BE910236	+	72.00	116.69	3.0e+03	1747	1	BE910236	601503560Fl NIH_MGC_7	gb_est2:BE618790	-	71.50	124.91	1.0e+03	786	1	BE618790	601462421Fl NIH_MGC
gb_est2:BE426162	+	71.50	134.13	317.73	354	1	BF426162	6194406 y1 Gm-cl047 GI	gb_est2:BE673868	-	71.50	124.72	1.1e+03	799	1	BE673868	602484387Fl NIH_MGC
gb_est1:BE813064	-	71.50	133.62	339.26	370	1	BE813064	CM2.AN007-8-090800-321	gb_est2:BM110188	-	71.50	124.59	1.1e+03	808	1	BM110188	EST557724 potato ro
gb_est1:BB809162	+	71.50	133.28	354.93	381	1	BB809162	BB809162 RIKEN full-1e	gb_est2:BE708981	-	71.50	124.42	1.1e+03	820	1	BE708981	602671529Fl NIH_MGC
gb_est1:BA435711	+	71.50	132.38	397.91	412	1	BI435711	EST38472 P. infestans	gb_est2:BI148029	-	71.50	124.22	1.1e+03	834	1	BI148029	BA148029 unpublilhe
gb_est1:AM011731	+	71.50	132.35	399.34	431	1	AM011731	EST722553 Schistosoma	gb_gss:CN60202CF	-	71.50	124.15	1.1e+03	839	1	AL190968	Tetradon nigroviridis
gb_gss:BM106426	+	71.50	131.85	425.42	431	1	BM106426	510154 MARC 3BOV Bos t	gb_est2:BF274600	-	71.50	124.12	1.1e+03	841	1	BF274600	GA_EB002110Fl Goss
gb_gss:AC059336	-	71.50	131.62	438.66	440	1	AC059336	HS_5062.B2.H02.SP6E RH	gb_est2:BG829412	+	71.50	124.06	1.2e+03	846	1	BG829412	602752512Fl NIH_MGC
gb_est2:BG062545	+	71.50	131.59	440.14	441	1	BG062545	TaLc1147G1DR TaLc1 Tr	gb_est2:BE570511	-	71.50	124.06	1.2e+03	846	1	BE570511	601328631Fl NCI CGA
gb_gss:BM62412	-	71.50	130.61	499.07	480	1	BE356408	DGL_125.A04.B1.A002.Da	gb_est2:BF678917	-	71.50	124.02	1.2e+03	849	1	BF678917	HVSMEN0006014F Hord
gb_est2:BE578943	+	71.50	130.05	536.52	504	1	BG578943	dF02c09.y1 Wellcome CH	gb_est2:BF678917	-	71.50	123.97	1.3e+03	898	1	BF678917	602117535Fl Soares
gb_est1:AV921946	+	71.50	129.89	547.61	511	1	AV921946	AV921946 K. Scto unpubl	gb_est1:AL557984	-	71.50	123.26	1.3e+03	906	1	AL557984	AL557984 LTI.NFL008
gb_gss:AC082049	+	71.50	129.51	574.84	528	1	AO822049	HS_5483.AL.D02.T7A RQC	gb_est1:AL530605	-	71.50	123.05	1.3e+03	923	1	AL530605	AL530605 LTI.NFL003
gb_est2:BF076642	+	71.50	129.46	578.07	530	1	BF076642	226249 MARC 2BOV Bos t	gb_est1:AM667230	-	71.50	122.99	1.3e+03	928	1	AM667230	GA_Ea0007015 Gossy
gb_est1:AL120787	-	71.50	129.25	594.32	540	1	AL120787	unpubd07.x1 Stratagene	gb_gss:CN6050MTD	-	71.50	122.91	1.3e+03	934	1	AL458602	Tetradon nigroviridis
gb_est2:BE918075	-	71.50	129.18	595.95	541	1	BE918075	OVL_1.F03.B1.A002.Ovar	gb_est2:BI783455	-	71.50	122.88	1.3e+03	937	1	BI783455	603084616Fl NIH_MGC
gb_est2:BB355170	+	71.50	129.08	609.43	543	1	BE355170	DGL_39.G08.B1.A002.Dar	gb_gss:BM169347	-	71.50	122.53	1.4e+03	965	1	BM169347	602944110Fl NIH_MGC
gb_est2:BI4443181	+	71.50	129.08	607.43	548	1	BI4443181	daa99h08.y3 Wellcome G	gb_gss:BM168164	-	71.50	122.39	1.4e+03	969	1	BM168164	VS_Sba0002422F Grap
gb_est2:BF076061	+	71.50	129.04	610.72	550	1	BF076061	225459 MARC 2BOV Bos t	gb_est2:AL523474	-	71.50	122.39	1.5e+03	977	1	AL523474	AL523474 LTI.NFL003
gb_gss:BE076061	-	71.50	128.97	615.66	553	1	BE675722	PIL_29.G08.B2.A002.Pad	gb_est2:BG257957	-	71.50	122.07	1.5e+03	996	1	BG257957	602379415Fl NIH_MGC
em_estP1:BE365722	-	71.50	128.97	615.66	553	1	BE675722	PIL_29.G08.B2.A002.Pad	gb_est2:BE675957	-	71.50	122.07	1.5e+03	1005	1	BE675957	601334727Fl NIH_MGC
gb_est1:AL1833788	-	71.50	128.89	622.28	557	1	AL1833788	605095009.x2 G05 - Pnd	gb_est2:BI446668	-	71.50	122.87	1.5e+03	1022	1	BI446668	602913684Fl NCI CG
gb_gss:AC0625722	-	71.50	128.79	630.58	562	1	AC0625722	CITRBI-EL-265415.FP CIT	gb_est2:BM444947	-	71.50	121.66	1.6e+03	1041	1	BM444947	AGENCOURT.6387638
gb_est2:BI4443181	+	71.50	128.74	633.91	564	1	AM4443181	RCO-MT0003-140300-031	gb_hlc:AC000405	-	71.50	121.61	1.6e+03	1045	1	AC000405	Mus musculus 18 da
gb_est2:BI4443181	+	71.50	128.74	633.91	564	1	AM4443181	RCO-MT0003-140300-031	gb_gss:CN6044LP	-	71.50	121.24	1.7e+03	1079	1	AL311426	Tetradon nigroviridis
gb_est1:AV922060	+	71.50	128.66	640.59	568	1	AV922060	K. Scto unpubl	gb_est2:BE421046	-	71.50	120.72	1.8e+03	1129	1	BE421046	HMW005.C03 IREC HW
gb_est2:BG672167	-	71.50	128.58	647.29	572	1	BI680767	sal03B04.y1 Gm-cl050 G	gb_est2:BG029814	-	71.50	120.52	1.8e+03	1149	1	BG029814	602296548Fl NIH_MGC
gb_est2:BG672167	-	71.50	128.58	652.33	575	1	BG672167	PMO-BT0587-110200-001	gb_hlc:AC009754	-	71.50	119.03	2.2e+03	1307	1	AC009754	Mus musculus adult
gb_est2:BI444671	-	71.50	128.48	655.70	577	1	BI444671	daa74605.y1 NICHD XGC	gb_est2:BE011511	-	71.50	117.06	2.8e+03	1550	1	BE011511	602282144Fl NIH_MG
gb_est1:AL502082	+	71.50	128.30	670.92	586	1	AL503082	AL503082 Hordum vulg	gb_hlc:AF161316	-	71.50	113.59	4.4e+03	2091	1	AF161316	Homo sapiens HSPc1
gb_gss:BM658495	-	71.50	128.28	672.62	587	1	BM658495	1000201B02.x4 1000 - U	gb_est2:BI027953	-	71.00	138.05	192.29	231	1	BI027953	IL5-MT0259-15H2001-3
gb_gss:BM658495	-	71.50	128.28	672.62	587	1	BM658495	1000201B02.x4 1000 - U	gb_est2:BG535480	-	71.00	135.58	263.94	286	1	BG535480	602653163Fl NIH_MGC
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87	288.94	304	1	BM523777	PM2-HTQ353-281299-0
gb_est2:BI721504	+	71.50	128.03	694.83	600	1	BI788031	3214-63 Mouse E14.5 rd	gb_est1:BM523777	-	71.00	134.87					

qb_est1.AM188002	+	71.00	128.34	660.45	535	BM188002	sa188010.y1	Gm-cl0174
qb_est1.AM188006	+	71.00	128.08	690.15	547	BM188003	sa188010.y1	Gm-cl0174
qb_est1.AW65532	+	71.00	128.04	694.19	549	AW65532	da73605.y1	Harland Homo s7
qb_est1.AA393225	+	71.00	127.83	713.03	559	AA392265	LD11881.5pr1me	LD DrosG16
qb_est1.BM337289	+	71.00	127.54	739.67	573	BM337299	MESTR205-H03.T3	ISUMS v5
qb_est1.BG641822	+	71.00	127.44	749.26	578	BG641822	pg1.c.pK002.d14	Chickadee
qb_est1.B3321218	+	71.00	127.30	762.76	585	B3321218	602961d466F1	NC1-CGAP
qb_gms.h2394035	-	71.00	127.30	762.76	585	AZ334035	IM0157J20 Mouse	10kb
qb_est1.B1588390	+	71.00	127.28	764.56	586	B1588390	RH30611.5pr1me	RH DrosG16
qb_est1.BE347372	+	71.00	127.24	768.56	588	BE347372	sp33905.y1	Gm-cl043
qb_est1.AA016910	+	71.00	127.20	772.44	590	AA016910	mh34f11.r1	Soares mouse

BASE COUNT	208	a	253	c	271	g	139	t
ORIGIN								
alignment_scores:								
Quality:	97.50				Length: 278			
Ratio:	0.806				Gaps: 15			
Percent Similarity:	43.525				Percent Identity: 24.101			

seq_name: gb_gss:CNS04NYY

15 CTCGAGATACAGACAGCCCCCTTCAGGCCAGCCAGCTCTCCCTA 64
18 gSerGlyLeuMetProArgLy.....HisasnGlyTrpPheAspA 33
65 CACC...GCCACCATGCCATTCGGTAACACCCACAACAAGTTC..... 104
33 rGlyThrGlnMetAsnLeuAsnLeuTyraSnaHisAlaArgGlyThrGln 49
105AGGCTGAATTACAAAGCCTGAGGAG..... 128
50 ThrGlyPheValArgTyraSnaSnaGlyTyraValSerThrSerLeuSerLe 66
129GAGTACCCCGAC.....CTCAGACCA 148
66 uArgSerAlaHisLeuAlaGlyLeuSerLeuSerGlyTyrrSerThrT 83
149 ACATAAACAACACATGGGCCAAGTACTGACCTTGAACTCTTCAAGAAGC 198
83 yTyrrLeuTyraValLeuAlaThrAlaProAsnMetPheAsnValAsnAsp 99
199 TCGCG.....GACAAGAGACTCCATTCGGCTTCACTGATGACGAT 239
100 ValLeu.....GlyValTyrrSerPro...HisProTyrrGlnGlnValA 113
240 GTCATCCACAGACAGAGAGTGGACAACCCAGGTGACCCCTTCATCATGACCT 289
113 lSerAlaLeuGlyGlyLeuProTyrrSerGlnLeuTyrrGlyTrpTyrrArgV 130
290 GGGCTGGCGTGGCTGGATGAGAGATCC.....TACGAGATTTCACAG 333
130 AlaSnaPheGlyValIleAspGluArgLeuHisArgAsnArgGlyLutyrArg 146
334 AACCTCTTTGACCCCATTCGATCGGATCGGACGCGGGCTTACAAACCCACT 383
147 AsparGlyTyrrArgAsnLeuAsn..... 154
384 GACAAACACAAAGACGTGACTCAACCATGAAGAACCTCAAGGTGGAGACGA 433
155IleAlaProAlaG 159
434 CCGGACCCCTAATCACTAGCTGCTCAGACGCCGCTCGGACTCGCGCGACGCA 483
159 lAspGlyTyrrArgLeuAlaGlyPheProProAspHisGlnAlaTrpArg 175
484 TCAAGGCTACACGTTG.....CCCCCA..... 506
176 GluGluProTrpIleHisHisAlaProGlnGly..... 186
507CACTGTCCCGTGGGAGCGCGCGGTGGAGA 538
187CysGlyAsnSer.....SerArgT 193
539 AGCTCTCTGTGGAGCTCTCAACAGCTGAGGCCGAGGCTCAAGAGGAGA 588
193 hrIleThrGlyAspThrCysAsnGluGlnThrGlnAsnLeuSerThrIle 209
589 CTAAACCTGTGAGACATGAGGGGAAGACACAGCAGATCATGGATGACA 638
210 TyrrLeuArgGlyTrpGlnSerLySnaValysaArgGlnIlePheSerAspTy 226
639 TTTCTGTGCAAAAGCGGTGTCCTGTCGGCGGCGGTCAAGGATGGCCGATGG 688
226 rGlnSerGluValAspIleTyrrAsnArgIleArg 237
689 GCCAAGCCGGGTCTGGGCAAGGACAGACTTGG 722

sequence.
 accession AL299347.1 GI:8038488
 version GSI: genome survey sequence.
 keywords Tetradodon nigroviridis.
 source Tetradodon nigroviridis
 organism Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthomorpha; Acanthopterygii; Percomorpha; Tetradontiformes; Tetradontidae; Tetradon.
 1 (bases 1 to 812)
 Roest-Crolius,H., Jalllon,O., Dasilva,C., Fizesma,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradodon nigroviridis
 Unpublished
 2 (bases 1 to 812)
 Roest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizesma,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using Tetradodon nigroviridis DNA sequence
 Unpublished
 3 (bases 1 to 812)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetradon>.
 Location/Qualifiers
 1. 812
 /organism="Tetradodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="123D09"
 /clone.lib="G"
 /note="Genoscope sequence ID : C08G123C05Lp1-end : 77"
 BASE COUNT 154 a 260 c 231 g 164 t 3 others
 ORIGIN

alignment_scores:
 Quality: 93.50 Length: 250
 Ratio: 0.766 Gaps: 14
 Percent Similarity: 48.800 Percent Identity: 24.000

Alignment block:
 US-09-528-682-1 x CNS04NXY ..

Align seg 1/1 to: CNS04NXY from: 1 to: 812

```

5 LeuTYATGAlaAspSerArgProPasp.....GlutIlely 17
||| ||| ::||| ||| ||| ||| ||| ||| ||| ||| |||
45 GCGCTAAGGAGCAGCTGCTCAGTTGACACAGCGCTGAGTGA 94
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17 SARSGerGlyGlyLeuMetProArg.....GlyHSAasnGly 31
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 A.....GGAATGTTACCGCAGAACCGCGTGGAGCAGACTGAC 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
31 heAspArgGlyThrGlnMetAsnIleAsnLeuTyAspHisAlaArgGly 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 .....AGCGCTGGCAGACGAGCTGAGGATCAGTCAAGTCA 173
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 ThrGlnThr...GlyPheValArgTyAspArgGlyTyValSerThr 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
174 GGGCCACATTCAGTGTGATGATGAGCTCAGTATGATACAGCTA 223
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 rIeuSerLeuArgSerAlaHisLeuAlaGlyInsSerIleLeuSerGly 80
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 ACCCAACATGAGCCGACGCTCTGTGAATGACACTGCCACTCCGAA 273
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

80 yISeThrTyTrpIleTyValIleAlaThrAlaProAsnMetPheAsn 96
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
274 CCAAGC.....AGAGCTCCACACCC.....TTC 296
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 ValAsnAspValLeuGlyValTySerProHisProTyArgGlnGluVal 113
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 ACCTGCTCCACAGTACTCTGATTC.....CAGATCAGGTGCT 337
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 lSerAlaLeuGlyIleProTySerGlnIleTyGlyTrpTyArgVal 130
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 CTCCCTCCAGAGGAGCGTGGG...CGCACGTGAGCGCTTCACCGTG 384
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 AlaSnPheGlyValIleAspArgLeuHis...ArgAsnArgGlu... 144
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 CCGCTATGGGAGATTGACACCGCTCCATCTTATGAGCGAGAGAC 434
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 .....TyArgAspArgTyTyArgAsnLe 153
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 GCACAGCGGTGGCCCTTCCACGTGCGAGGAGCGCGCGTCAAC 484
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153 uAsnIleAlaProAla..... 158
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
485 GGGAGCGCTGCGCGGGCGGTCCGAGAGCTCCACCTTGCGCT 534
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 .....GluAspGlyTyArgLeuAlaGlyPhePro 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
535 CCGACCTGCACCTTGACCTCCAGACTT..TTCAGGACATGGGGCTCCA 583
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169 ProAspHisGlnAlaTrpArgGlnIleProTyP..... 179
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 CCGGAC...CTTGATGAGACTTGTGAGCTGGGATGAGAACTCATCTCC 630
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 .lIeHisHisAlaProGlnGlyCysGlyAsnSerSerArgTrpHisThrG 196
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
631 TCGAGACGAGTACAGCTGTCATGTTCAACTTTTCCGTTCCTCCG 680
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 lYAspTrpCysAsnGlnGluThrGlnAsnLeuSerHisIleTyLeuArg 212
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
681 CAGGTGTTGCTGCTTCAGCATTTGACGAGGCTCTCCGACGCGAGCGC 730
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: gb_est2:BG171693

seq documentation block:
 LOCUS BG171693 939 bp mRNA linear EST_06-FEB-2001
 DEFINITION 602322217F1 NIH_MGC_89 Homo sapiens cDNA IMAGE:4425124 5',
 mRNA sequence.
 ACCESSION BG171693
 VERSION BG171693.1 GI:12678396
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMLN at:
<http://image.llnl.gov>
 Plate: L1AM10170 row: b column: 05
 High quality sequence stop: 695.
 Location/Qualifiers
 1. 939
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4425124"

/clone_lib="NIH_MGC_89"
 /tissue_type="hypenephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 246 a 231 c 229 g 233 t
 ORIGIN

alignment_scores:

Quality: 93.00 Length: 261
 Ratio: 0.838 Gaps: 13
 Percent Similarity: 42.529 Percent Identity: 19.540

alignment_block:

US-09-528-682-1 x BG171693 ..

Align seg 1/1 to: BG171693 from: 1 to: 939

```

4  AAGLeuTyrArgAlaAspSerArgProAspGluIleLysArgSerG1 20
   ::::: |||::: ||| ::::: |||
227 AAGTCATGCTGCTGAGATATTCCTGAGAAATCCCTGAAATGCTGTA 276
   ::::: |||::: ||| ::::: |||
20  YGLeuMetProArgGlyHis.....AsnGluTyrPheAspArgGlyT 35
   ::::: |||::: ||| ::::: |||
277 TCCAAAATATGACAAAGACCCTTGTGAAAGTATCTACCAAGGCCAA 326
   ::::: |||::: ||| ::::: |||
35  hrcIn.....MetAsnIleAsnLeuTyrAspHisAlaArgGly 47
   ::::: |||::: ||| ::::: |||
327 AGAAGCAGACAGAGCCATGAGACTT.....GATGCCCTTAAGGA 367
   ::::: |||::: ||| ::::: |||
48  ThGlnThiGlyPheValArgTyrAspAspGlyTyrValSerThrSerle 64
   ||| ||| ::::: |||::: |||
368 ACT.....GGATATATCAAGACTGAGT 390
   |||
64  uSerLeuArgSerAlaHisLeuAlaGlyInsIleLeuSerGlyTyrS 81
   |
391 G..... 391
81  erThrTyrTyrIleTyrValIleAlaThAlaProAsnMetPheAsnVal 97
   ||| ||| ::::: |||::: |||
392 .....ATTCTGTGCTGAGTTCACCCCTTCTAGACTTCAGACC 430
   ::::: |||::: ||| ::::: |||
98  AsnAspValLeuGlyValTyrSerProHisProTyrGluInsIleValSe 114
   ::::: |||::: ||| ::::: |||
431 ACAGACACACCTGCTCCCATGTCCTGAGAGAGTTGACGAGTGTCTCG 480
   ::::: |||::: ||| ::::: |||
114  rAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly..... 126
   ::::: |||::: ||| ::::: |||
481 GATAGTGGGCTGTGTAATTCACAGTATGATGACAGTATAGACGA 530
   ::::: |||::: ||| ::::: |||
127 .....TyrTyrArgValAsnPheGly..... 133
   ::::: |||::: ||| ::::: |||
531 TGAATTTTTCATCTTCTGCGCAGAGTTTCTTCTCATCTGTGAT 580
   ::::: |||::: ||| ::::: |||
133 ..... 133
581 CCCTCTCTACTCTGTCTTCATCTCCTGTGTTTTCAGGAATGAAA 630
   ::::: |||::: ||| ::::: |||
134 .....ValIleAspGluArgLeuHis 141
   ::::: |||::: ||| ::::: |||
631 CGAAAGCGACAAATTCGTCGACAACTGTGATAGCAAGATTTTTC 680
   ::::: |||::: ||| ::::: |||
141 rGAsnArgGlu.....TyrArgAspArgTyrTyrArgAsnLeuAsn 154
   ||| ||| ::::: |||::: |||
681 CTAACTCAGAAACATCAAGTTACTTGAAAGGATCATGATCTTACGA 720
   ::::: |||::: ||| ::::: |||
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
   ::::: |||::: ||| ::::: |||
731 GTA.....ACTGACAGGCTCTCCGACAAAGGGGGTTCCCGGAAACA 774

```

```

171 s.....GlnAlaTrpArgGluInsuProTyr..... 179
   |||::: |||::: |||
775 TCGCATTTCCCAAGTTTGAGAGATGAGTCTGGGAAAGAAATTCCTCT 824
   ::::: |||::: ||| ::::: |||
180 .....IleHisAlaProGlnGlyCysGlyAsnSerSerArgThr 193
   ::::: |||::: ||| ::::: |||
825 CCAGTTTGTAGTTCGCCCTCAGAGG...GGCAACACGCTTGCCACG 871
   ::::: |||::: ||| ::::: |||
194 IleThr...GlyAspThrCysAsnGluInsuThr 203
   ||| |||::: |||::: |||
872 CATACCCCTGGGACAAATTTCTGTAGAGACC 904

```

seq_name: gb_gss:A2698180

seq_documentation_block:

LOCUS A2698180 714 bp DNA linear GSS 24-JAN-2001
 DEFINITION RPCI-23-223D11.TU RPCI-23 Mus musculus genomic clone RPCI-23-223D11
 , DNA sequence.

ACCESSION A2698180
 VERSION A2698180.1 GI:12415007

KEYWORDS
 GSS.

SOURCE
 house mouse.

ORGANISM
 Mus musculus

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 other GSSs: RPCI-23-223D11.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhaod@igf.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tigr/bac_ends/mouse/bac_end_intro.html
 Plate: 223 row: D column: 11
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source

Location/Qualifiers
 1..714
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-223D11"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: kidney/Brain; Vector: pBAC3.6; Site: 1:
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI methylase. Size
 selected DNA was cloned into the pBAC3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 191 a 179 c 181 g 163 t
 ORIGIN

alignment_scores:

Quality: 90.50 Length: 80
 Ratio: 1.708 Gaps: 6
 Percent Similarity: 66.250 Percent Identity: 37.500


```

was primed with a Not I - oligo(dT) primer [5',
TCGTACCAATCTGACAGTGGGAGGCGGCCGCTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CAAGATTCGGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT731 vector.
Library constructed by Bob Barstead."

```

BASE COUNT	128 a	148 c	140 g	124 t	3 others
------------	-------	-------	-------	-------	----------

```

ORIGIN

```

alignment_scores:					
Quality:	89.00		Length:	117	
Ratio:	1.618		Gaps:	8	
Percent Similarity:	47.009		Percent Identity:	29.060	

```

alignment_block:
US-09-528-682-1 x AA473437 ..

```

Align seg 1/1	to: AA473437	from: 1	to: 543
---------------	--------------	---------	---------

```

125 TYRGTYTTPYRARGVALAsnPhgeGLV.....IleAsp...GLuAR 138
      ||| ||||| : : : : : ||| ||| ||| |||
147 TATGAGTGTTCATCTACGAGTATGAGGCCCTCACATCGATGGGAGGA 196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 gleuHISARGAsnARGLuTYRARGAspARYTYR..... 149
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197 ATACATTCCTTTTAAOCAGATATGACAGCAAAATATATACCTCTTGTCACAG 246
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 ..... TYRARGAsnLeuAsnIleAla 156
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 TAGCCAGACTACTGCTCGACAGACCAATACCTGGAACGTAGATGCACATTA 296
157 PROAlaGLu.....AspGLYTYRARGLeuAlaGLYPhePROProAs 170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 CCAAGAGAACTTGGGGCAATTGGCTTGGTCATCTTGGGCTTT..CTTTC 343
170 pHisGLnAlaTPARgGLuGLu.....PROT 179
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 CAACCAAAATTTGGAAACAGGAACAGACGAGAACTCCGAGATACATCCCA 393
179 pRLeHISHisAlaPROGLnGLY.....CyGLYAsnSerSerARg 192
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
394 GTCTCAGATGTCTCGACAGAGGTGGGCTTTGTCTAATTTCCAGCTC 443
444 CTTGAGAAAGAGAGATGTGANCGGGGAAGAGAAGACAGAAAT...TCTACACT 490
209 e 209
:
491 T 491

```

seq_name: gb_est2:BI910238	
----------------------------	--

seq_documentation_block:					
LOCUS	BI910238	1147 bp	mRNA	linear	EST 16-Oct-2001
DEFINITION	603069452F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5218236 5',				
ACCESSION	BI910238				
VERSION	BI910238.1 GI:16173620				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 1147)				
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabs-remail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	CDNA Library Preparation: Life Technologies, Inc.				

Ratio: 1.630 Gaps: 6
Percent Similarity: 49.091 Percent Identity: 29.091

alignment_block:

US-09-528-682-1 x B1725716/rev ..

Align seg 1/1 to reverse of: B1725716 from: 1 to: 558

```

98 AsnAspValLeuGlyValTyrSerProHis.....Pr 108
   ::::::::::::::::::::
530 GAGAGTCTCTGGCTGTGTACAGCCCGCCCATTTAGTCATCCACA 481
   ::::::::::::::::::::
108 CTYrGluGlnGluValSerAlaLeuGly GlyLeuProTyrSerGlnIle 124
   ::::::::::::::::::::
480 GCACCGACGAGATCCAGCGGCTGTGGCGGCTTGCATCTTCTTCTC 431
   ::::::::::::::::::::
125 TYrGlyTyrTyrArgValAsnProGlyValIleAspGluArgLeuHisAr 141
   ::::::::::::::::::::
430 CAT.....TTCCGCTCGCGCTATCCGCTAGCTGACTGGGCACTT.... 392
   ::::::::::::::::::::
141 GAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProA 158
   ::::::::::::::::::::
391 .....GTCATCATTTGCCCG 376
   ::::::::::::::::::::
158 IaGlu.AspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTr 174
   ::::::::::::::::::::
375 CAGAGCTCCGCGGTACGCCAGCTCGTCAAGCGCCCTCCGCGCGCTGGAG 326
   ::::::::::::::::::::
174 PArgGluGlnProTyrIleHisHisAlaProGlnGlyCysGlyAsnSer 191
   ::::::::::::::::::::
325 ACCGCGAG.....ACAA 315
   ::::::::::::::::::::
191 eArgThrIleThrGlyAspThrCys 199
   ::::::::::::::::::::
314 GCCCAACGATAGGAGCGATCCCTGC 289
   ::::::::::::::::::::

```

seq_name: gb_gss:AQ159897

seq_documentation_block:

LOCUS AQ159897 698 bp DNA linear GSS 09-SEP-1998
DEFINITION mgx0002M14r CUG1 Rice Blast BAC Library Magnaporthe grisea genomic
clone mgx0002M14r, DNA sequence.

ACCESSION AQ159897.1 GI:3556886

VERSION AQ159897.1 GI:3556886

KEYWORDS GSS.

SOURCE Magnaporthe grisea.

ORGANISM Magnaporthe grisea.

REFERENCE 1 (bases 1 to 698)

AUTHORS Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
Phillips,K., Sasinowski,M., Wing,R.A. and Dean,R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdeane@clemson.edu
Seq primer: GGAACAGCTAGACCATG
Class: BAC ends
High quality sequence stop: 417.

FEATURES Location/Qualifiers

1..698

1..698

1..698

1..698

1..698

1..698

/lab_host="E. coli DH10B"
/note="Vector: pBACNICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25x genome coverage. High density colony filters
are available upon request."

BASE COUNT 133 a 197 c 194 g 174 t
ORIGIN

alignment_scores:
Quality: 88.00 Length: 104
Ratio: 1.443 Gaps: 5
Percent Similarity: 58.654 Percent Identity: 27.885

alignment_block:
US-09-528-682-1 x AQ159897/rev ..

Align seg 1/1 to reverse of: AQ159897 from: 1 to: 698

```

31 PheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgI 47
   ::::::::::::::::::::
695 TTTCAGACCTGGACATGACATTCATCAGATGTTATGATCACCACAGCA 646
   ::::::::::::::::::::
47 YThrGlnThrGlyPheValArgTyrAspArgGly...TyrValSerThr 63
   ::::::::::::::::::::
645 AAATGAA.....CGTAGACAAAGAGCCGCTTATTCGACTT 608
   ::::::::::::::::::::
63 eLeu.....SerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 77
   ::::::::::::::::::::
607 CGGCGAACAAGGCAATCTCCCGCAGCATGTCCAGATAGAAACCGGGCC 558
   ::::::::::::::::::::
78 SerGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsn 94
   ::::::::::::::::::::
557 AAGTAGTAGTACGATATTCATGATGACAGCGCCGATCCGCGCAACTT 508
   ::::::::::::::::::::
94 tPheAsnVal.....AsnAspValLeuGlyValTyrSerProH 107
   ::::::::::::::::::::
507 TATCGACATTCGCGCAGTGTGACAGCATCGGAGGAGCATCAT..T 461
   ::::::::::::::::::::
107 IsProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyrSerGln 123
   ::::::::::::::::::::
460 GGCCTTGGAGAGAGAGATTGCTGCCAAGCGCAAAATCCCTGCTCCAGC 411
   ::::::::::::::::::::
124 IleTyrGlyTyr 127
   ::::::::::::::::::::
410 GTCGTGCGCTAC 399
   ::::::::::::::::::::

```

seq_name: gb_hlc:BC012102

seq_documentation_block:

LOCUS BC012102 1982 bp mRNA linear HTC 06-AUG-2001
DEFINITION Homo sapiens, similar to protein tyrosine phosphatase,
receptor-type, F, clone IMAGE:4661274, mRNA.

ACCESSION BC012102.1 GI:15082377

VERSION BC012102.1 GI:15082377

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1982)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

/lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt:
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 309 a 379 c 404 g 333 t others

alignment_scores:
 Quality: 87.00 Length: 306
 Ratio: 0.690 Gaps: 16
 Percent Similarity: 41.176 Percent Identity: 20.261

alignment_block:
 US-09-528-682-1 x B1663851/rev ..

Align seg 1/1 to reverse of: B1663851 from: 1 to: 1426

```

2  G1YAspArgLeuTyrrArgAlaAspSer..... 10
876 GGGTCCAAACTGTGTAAAGGAGCCGCCCTTTCGTGATCGCTAAACT 827
11 .....ArgProAspGluLeuYsArgSerGlyLeuMetP 24
826 GTTTCGCTTTCGGCCCCCT.....CCCCGGGAGACCTCTTCTCC 786
24  rArGclYhIsAnGluTyrrPheAspArgGlyThrGlnMetAnLeAsn 40
785 CAAGGGGTAT..... 775
41  LeuTyAspRHisAlaArgGlyThrGlnThrGlyPheValArgTyrrAsp.. 56
774 .....CACACTTCAGGGGGGGGCTCTGTTCACGGCTTAAACGG 734
57 .....AspGlyT 59
733 GGGACAGGGGTCCCACTTCGCACACAGTCCTGGCCCAAGAGATGGTC 684
59  yrrValSerThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyIns 75
683 AATGTGCAAAATCATCTCTACGTCGTCTCTAGAACATCTCC 634
76  IleLeu.....Se 78
633 TTGTTTCCTGGGGCACAAGGAGGCTTTTCCCCATGATGGGCCA 584
78  rGlyTyrrSerThrTyrrIleTyrrValIleAlaThrAlaProAsn.... 93
583 TGGTTCCTCTGTCGTTACGGGGGATAGTGGATGGCCCTCGCCCA 534
93 ..... 93
533 GACCATTAACGATGCCAAATTCAGATGATCATGTACTAGACT 484
94 .....MetPheAsnValAsnAspValLeuGly ValTyrrSerP 106
483 TGCAGCGATGTTTCATTCACGTTGCACTCCATGGGCTATGTTTGGC 434
106 rHisProTyrrGluGln.....GluValSerAlaLeu... 116
433 CTCACAAAGACAGAAACCCATGTGTAATATGATCTCGCACCATCAT 384
117 .....GlyGlyLeuProTyrrSerGlnIleTyrrGly.....TrrTyrrAr 129
383 CGAGACCTGGGGCTGAGATATACTCTTCTTACGCTCAACAACGGGAG 334
129 9...ValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlu 145
333 TCCCGCTCGCTTTCATGCTACTACACGC..... 303
145 yrrArgAspArgTyrrTyrrArgAsnLeuAsnIleAlaProAlaGluAspGly 161

```

```

302 .....CAATCGCTAATTCTAGCTCTTCGACAGACAGTC 270
162 TyrrArgLeuAlaGlyPheProAspHisGlnAlaTrpArgGluGluP 178
269 AACAGATTCACAGCTCTCAGC.....CGCAGCTCTCC 238
178 otrPleHis.....HisAlaProGlnGlyCys.....G 188
237 GATTCTCCACCTTCATTCAGACATGACGGGACACATGCCACAGCTG 188
188 1YasnsrSerArgThrIleThrGlyAspThrCysAnGluGluThrGln 204
187 GCCTGAGAGGAGAACCTCGATGATGATGTGTGCTGGCCCTGCGGT 138
205 AsnLeuSerThrIleTyrrLeuArgGluTyrrGlnSerIlyValysArg 221
137 GCCCTGCGCTGACTGCACTGTCTGAGC...AGAATGACGACGAGAGAC 91
221 nIlePheSerAspTyrr 226
90 TGTAGACCGAGACTAC 75
seq_name: gb_est1:AA145382

```

seq_documentation_block:
 LOCUS AA145382 497 bp mRNA linear EST 11-FEB-1997
 DEFINITION mr78g1.r1 Strati gene mouse heart (#937316) Mus musculus cDNA clone
 IMAGE:603620 5', mRNA sequence.
 ACCESSION AA145382
 VERSION AA145382.1 GI:1714756
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 497)
 AUTHORs Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Scheinberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The Mashu-HHMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HHMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LINL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:369052
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 415.

FEATURES

source

1..497 Location/Qualifiers

```

/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:603620"
/clone_lib="Strati gene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site:1;  

  EcorI; Site_2: XhoI; Cloned unidirectionally. Primer:  

  Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.  

  Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'  

  adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'."

```

BASE COUNT 158 a 118 c 126 g 95 t


```

169 .....ProAspHisGlnAlaIlePrrpg 175
407 TGCACACTGCCCGCAGACC GGAGCCGGCAGGAGGCGCTGGCGG 450
seq_name: gb_gss:CNS03001

seq_documentation_block:
LOCUS      CNS03001                905 bp    DNA       linear   GSS 15-MAY-2000
DEFINITION Tetradon nigriviridis genome survey sequence T7 end of clone
            183N05 of library G from Tetradon nigriviridis, genomic survey
            sequence.
ACCESSION  AL221643
VERSION    AL221643
KEYWORDS   GSS; genome survey sequence.
SOURCE     Tetradon nigriviridis.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Tetraodontiformes; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontifformes;
            Tetraodonidae; Tetraodon.
REFERENCE  1 (bases 1 to 905)
AUTHORS   Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetradon nigriviridis
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 905)
AUTHORS   Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottler,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE     Human gene number estimate provided by genome wide analysis using
            Tetradon nigriviridis DNA sequence
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 905)
AUTHORS   Genoscope.
TITLE     Direct Submission
COMMENT    Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
            This sequence is a single read and was generated as part of a large
            scale clone-and-sequencing project of the Tetradon nigriviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.

FEATURES             Location/Qualifiers
     source           1..905
                     /organism="Tetradon nigriviridis"
                     /db_xref="taxon:99883"
                     /clone="183N05"
                     /clone_1bp="c"
                     /note="Genoscope sequence ID : C0Ng183CG03LP1-end : T7"
BASE COUNT          163 a      340 c      274 g      127 t              1 others
ORIGIN
alignment_scores:
    Quality:      85.50      Length:      201
    Ratio:        0.847      Gaps:        10
Percent Similarity: 50.249   Percent Identity: 23.881

alignment_block:
US-09-528-682-1 x CNS03001 ..

Align seg 1/1 to: CNS03001 from: 1 to: 905

13 ProAspHisGlnAlaIlePrrpgSerArgGlyLeuMetPro.ArgGlyHisAsnG 29
|||||::: ::||| ||||| |::: ||| ::||| ||| |
216 CGCGAGAGATGGACCGCGGTCTGGAGGAAGATGTCCCTGCACGGCACCGCG 265
29 LuTrpheAspArgGlyThrGlnMetAsnIleAsnLeuTyraSPHisaLa 45
:: ::::::||| :::: :::: ||| |||:::
266 AC...CTGTCGCCAGCGCAGACGCGCATGCTG.....CACTGG 306
46 ArgGlyThrGlnThrGlyPheValArgTYraSPASpGLTYryValSerTh 62
||||| ::::: ||| ||| ||| |||

```

```

307 CGGAGACTCTCGGCCCTCTTCCCAAGGCCCCCTG666GGCTCCCTCAT 356
62 rSerLeu.....SerLeuArgS 68
   :|||
   :|||
357 GGGCGTGGAGCAGCAGCAGCGGTCGCCGCGCTTCAGCACCATGCTCGG 406
68 eArLaHisLeuAlaGlyInSerLLeuSerGlyTygSerThrTyTyTy 84
   :|||
   :|||
407 CGGTGGGCTACAGCGGCGACTGCGCTCCAGCTCGGGGAGCAGCTC... 453
85 lLeTyValLlLeaThrAlaProAsnMetPheAsnValAsnAspValLe 101
454 .....ACCACTGAGCCCGCGTGCAGACCTTGAGCAGCAAGTT 491
101 uGlyValTygSerProHisProTygInGluInGluValSerAlaLeuGly 118
   :|||
   :|||
492 CCGACCCAGCAGCAGCCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAG 528
118 lYlLeProTygSerGlnLlLeTyGlyTTPTyArgValAsnPheGlyVal 134
529 ...CTGCCCGCTCAGCAGCGATCGGAGCTTCACCTG.....ATG 567
135 lLeAspGlyArgLeuHisArgAsnArgGluTyArgAspArgTyTyTy 151
568 CGGAGAGAGCGGGCGCAGCTGGGCGACCACTTACACCCCTCAGGGAA 617
151 gAsnLeu.....AsnLlLeAlaProAlaGluAspGlyTyTyArgL 164
618 GGACCTGGCGCATGTCAGAGCCTGTGCGCCCGCTCAGCGGGGCGGGCG 667
164 eAlaAlaGlyPheProProAspHis.GlnAlaTPArgGluInGluProT 180
668 TGGGCTCTCCCTGTCAGCGCTACGGCAGCGCTGGGCGAGCAGCCCAAGCG 717
180 eHisHisAlaProGlnGlyCysGlyAsnSerSerArgThrLleThyGly 196
718 AACGGCAGCC....AGATGCTCACGGGCGGCTACGAGCTCCACGGGG 760

seq_name: gb_est2:B1546037

seq_documentation_block:
LOCUS      B1546037          931 bp      mRNA      linear      EST_05-SEP-2001
DEFINITION 603188178B1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5259786 5',
            mRNA sequence.
ACCESSION  B1546037
VERSION    B1546037.1 GI:15433349
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 931)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-research.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shih-Wei
            Toshiyuki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/MLN at:
            http://image.llnl.gov
            Plate: LLM11634 row: k column: 19
            High quality sequence stop: 778.
            Location/Qualifiers
                1..931
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5259786"
                /clone_id="NIH_MGC_95"
FEATURES
source

```


415 CTTGACGCGCGGCTCGGACTCGGCTGAGCCCTTCGAGCGGATACCGGA 464
 189 AsnSerSerArgThr 193
 |||||
 465 TGGCTCTTCGAAAC 479

seq_name: gb_gss:A2935912

seq_documentation_block:
 LOCUS A2935912 1039 bp DNA linear GSS 21-DEC-2001
 DEFINITION ODS534 Oikopleura dioica Shotgun library Oikopleura dioica genomic
 clone 6534, DNA sequence.
 ACCESSION A2935912
 VERSION A2935912.1 GI:17976083
 KEYWORDS GSS.
 SOURCE Oikopleura dioica.
 ORGANISM Oikopleura dioica.
 Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
 Oikopleuridae; Oikopleura.
 1 (bases 1 to 1039)
 REFERENCE Seo,H.C., Kube,M., Edvardsen,R.B., Jensen,M.F., Beck,A., Spriet,E.,
 Gorský,G., Thompson,E., Lehrach,H., Reinhardt,R. and Chourrout,D.
 Miniature genome in the marine chordate Oikopleura dioica
 Science 294 (3551), 2506 (2001)
 TITLE JOURNAL
 COMMENT Sars International Centre for Molecular Marine Biology
 Bergen High Technology Centre, Thormøhlensgt. 55, N-5008 Bergen,
 Norway
 Tel: 55 58 4389
 Fax: 55 58 4305
 Email: hee-chan.seo@sars.uib.no
 Insert length: 1039 Std Error: 0.00
 Seq primer: 17 and V4 (ATCCACGCGTGGAGCTCT)
 Class: Shotgun.
 Location/Qualifiers
 1..1039
 /organism="Oikopleura dioica"
 /db_xref="taxon:34765"
 /clone="G534"
 /clone.lib="Oikopleura dioica Shotgun Library"
 /sex="male"
 /note="Organ: sperm; Vector: pGBM-T Easy; The shotgun
 library was prepared using the EcoRV digested pGBM-T Easy
 vector (Promega) by Dr Hee-Chan Seo."
 BASE COUNT 291 a 199 c 223 g 326 t
 ORIGIN

alignment_scores:
 Quality: 83.50 Length: 249
 Ratio: 0.732 Gaps: 15
 Percent Similarity: 45.783 Percent Identity: 21.285

alignment_block:
 US-09-528-682-1 x A2935912/rev ..

Align seg 1/1 to reverse of: A2935912 from: 1 to: 1039

21 GYLeuMeProArGlyHISaNGluTYrPhaSPaRgLYThrGlnMe 37
 |||:||||| :|||:|||||:||||| :|||:|
 664 GGAATATATG.....AACAAATGTCAAATTCGATCAGGAGACAGT 624
 37 tAsnIleasNleuTYrAspHisAlaArgGlyThrGlnThrGlyPhe.... 52
 : :|||:||||| :|||:||||| :|||:||||| :|||:|
 623 G.....CCAAGTGTAGAAAGTTATGAAA 601
 53 ..ValArgTYrAspSPaRgLYrValSerThrSer..... 63
 ||| :|||:||||| :|||:||||| :|||:||||| :|||:|
 600 GTGTCTCTTCGATAAAGTTACACATTGAAGCAACTGACAGCCATAT 551
 64 LeuSerIleuArgSerAlaHis..... 70
 |||:| :|||:||||| :|||:||||| :|||:||||| :|||:|
 550 TTACAGGAGAGCTTGGCTATTCTTCATCAAAAGTCTTCATGCCCTGAA 501

71LeuAlaGlyInSerIleLeuSerGlyTyrSer 82
 ||| :|||:||||| :|||:||||| :|||:||||| :|||:|
 500 AAATGCTGCTCAAGAGCTGGAAGAGAAAAAATATCTCAGATCATGAA 451
 82 hTYrTYrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsn 98
 || :|||:||||| :|||:||||| :|||:||||| :|||:|
 450 CAATGCAT...TGGATTCTGTTGGCGATGCAATTTCCACTTCATTAATTCAG 404
 99 AspValIleuGlyValTYrSerProHisProTYrGluInGluValSerAl 115
 :|||:||||| :|||:||||| :|||:||||| :|||:|
 403 CCGAAATGTTCCTCATTTTCCGATCATGTCACCAT..... 365
 115 aleuGlyGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsn. 131
 |||:||||| :|||:||||| :|||:||||| :|||:|
 364TATAAGCAGCTTTATGGCTTACCGCTTCAGC 331
 132PheGlyValIleAspGluArgLeuHis..... 140
 |||:||||| :|||:||||| :|||:||||| :|||:|
 330 AGCATAAAAAATATCATTCGCAATTTATGAAGTTTCGAGTTTACAATGAC 281
 141ArgAsnArgGluTYrArgAspArgTYrArg... 151
 |||:||||| :|||:||||| :|||:||||| :|||:|
 280 CTCGTTGTAAAGACCATACAGACTACGACTATATTATTATTACGAAAA 231
 152AsnLeuAsnIleAlaProAlaGluAspGlyTYrArg 163
 |||:||||| :|||:||||| :|||:||||| :|||:|
 230 CCATCTAGTATGAAATTTAGAAATTTCTTCCTCAATTAATTTATTCAGA 181
 164 LeuAlaGlyPheProAspHisGlnAlaThrPargGluInGluProTyrP 180
 |||:||||| :|||:||||| :|||:||||| :|||:|
 180 TTA.....CTCATATTATTCGTCACAGTACGTCACCAAGTCCCT 143
 180 eHisHisAlaProGlnGlyCys.GlyAsn.....Ser 190
 |||:||||| :|||:||||| :|||:||||| :|||:|
 142 TCACCTGCATCCGAAAGATTGCACAAACAGATTTCGAAGATTTCAAA 93
 191 SerArgThrIleThrArgAspThrCysaNGluGluThrGlnAsnLeu 207
 :|||:||||| :|||:||||| :|||:||||| :|||:|
 92 TGCAAAACCTGCAGC...AATATTGGGACAGACACATTCAGAGACGCCA 46
 207 rThrIleTYrLeuArgGluTYrGlnSerLYsValLYsArgGln 221
 :|||:||||| :|||:||||| :|||:||||| :|||:|
 45 CGAAGTTTCTCGAAGACATCACTTCACAGCTTGAAGAGCCAG 3

seq_name: gb_gss:AQ955081

seq_documentation_block:
 LOCUS AQ955081 845 bp DNA linear GSS 27-JAN-2000
 DEFINITION nbe0080023f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
 clone OSJNB0080023f, DNA sequence.
 ACCESSION AQ955081
 VERSION AQ955081.1 GI:6778347
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoideae; Oryza.
 1 (bases 1 to 845)
 REFERENCE Wing,R.A. and Dean,R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
 TITLE JOURNAL
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GTAAACGACGCGCCAGTG
 Class: BAC ends
 High quality sequence start: 63


```

589 .....TATACACACACACCCATATATATACATGGC 620
90 hr1a1ProasmetPheasnValasnAspValLeuGlyValTyrSerPro 106
621 TATGTCCCTATCTATTCACAAACACTGGAAACTGGACCTCGCTCACCA 670
107 HisProTyrGluGluGlnValSerAlaLeuGlyLysIleProTyrSerG1 123
671 CAC.....GGAGCGGGAACACACATCA 693
123 nileTyGlyTyrPtyrArgValasnPheGlyValIleAspLysIleuH 140
694 AGCAGGACATCAGCACAATCCAAACCCACATCATACAAAGCCGAAACGC 743
140 IsArgasnArgGluTyrArgAspArg...TyrTyrArgasnLeuAsnIle 155
744 ACAACACCCACAGGACACATGACGACACACACACAGAGGACGACAG 793
156 AlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisG1 172
794 ACCCAACAGACAGGACGACAGGACAGACCAAGCCACGACGACAG 843
172 nAlaTPrArgGluGluProTPrIleHisAlaProGlnGlyCysGlyA 189
844 A.....CCACACAA..... 853
189 snSerSerArgThrIleThrglyAspThrCysasnGluGluThrGln 204
854 .....AGCAACACCCCGCGGGAACCCAGCAGACACACACACACAA 895

```

seq_name: gb_gss:AG031102

seq_documentation_block:

LOCUS AG031102 1068 bp DNA linear GSS 01-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-003L01.R, genomic survey sequence.
 ACCESSION AG031102
 VERSION AG031102.1 GI:16557975
 KEYWORDS GSS: GSS (genome survey sequence).
 SOURCE Pan troglodytes male lymphoblast DNA, clone_11b:PTB Chimpanzee Male
 BAC library clone:PTB-003L01.R.
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE BAC end sequences of library PTB

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1068)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Tsuchiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimpses@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. 1068
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="PTB-003L01.R"
 /sex="male"

FEATURES

source

/cell_type="lymphoblast"
 /clone_11b="PTB Chimpanzee Male BAC Library"
 BASE COUNT 170 a 426 c 171 g 176 t 125 others
 ORIGIN

alignment_scores:
 Quality: 83.00 Length: 130
 Ratio: 1.297 Gaps: 5
 Percent Similarity: 49.231 Percent Identity: 25.385

alignment_block:

US-09-528-682-1 x AG031102/rev ..

Align seg 1/1 to reverse of: AG031102 from: 1 to: 1068

```

59 TyrValSerThrSerLeuSerLeuArgSerLalaIleuAlaGlyGlnse 75
|||||.....:..||| :|||||.....
778 TACGTGCGTGTGGCTTCTCTACTCACACGCTCATCTGCGCCGACG 729
75 rIleLeuSerGlyTyrSerThr.....TyrTyrIleTyrV 87
|||:|||||.....
728 CGTACTTCATACACATCTACCGCGGCTCTCACAATTGATCTGTTTTT 679
|||:|||||.....
87 a11eAlaThr.....AlaProasmetPheasnValasn 98
:..:..||| :|||:..:..
678 TTCTACCTACTCAATATCATTTGCGATTCGCGTGTCTCATTTGACA 629
99 AspValLeuGlyValTyrSerProHisProTyrGluGlnGlnValSerAl 115
|||||.....
628 GCTAAGTTGGGCTGTGGCTGACCGCCCTCATCACAAGAGGGGGGACG 579
|||||.....
115 AleuGlyGlyIleProTyrSerGlnIleTyrGlyTPrTyrArgValasnP 132
|||||.....
578 GAGGGGGGCGC..... 569
132 heGlyValIleAspLysIleuHisArgasnArgGluTyrArgAspArg 148
:..:..||| :|||:..:..
568 .....GAGAGCGACAGACCAAGCGCGGAGTGGCGGACGCG 530
|||||.....
149 TyrTyrArgasnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAl 165
|||:|||||.....
529 TACACCCGACGATGAGA.....CGTCAAAATGTAGCTCTGAGAGCGC 486
|||||.....
165 aGlyPheProPro...AspHisGlnAlaTPrArgGluGlu 177
|||||.....
485 GCGGTGCCCGTCATRANGGAAGAGGGGTGGAGCGGAG 446
|||||.....

```

seq_name: gb_est2:BF429191

seq_documentation_block:

LOCUS BF429191 370 bp mRNA linear EST 29-NOV-2000
 DEFINITION WHE1709_E10_119ZS wheat heat stressed spike cDNA library Triticum
 aestivum cDNA clone WHE1709_E10_119, mRNA sequence.

ACCESSION BF429191 GI:11441115

VERSION BF429191.1

KEYWORDS EST.
 bread wheat.

SOURCE Triticum aestivum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 1 (bases 1 to 370)

REFERENCE Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
 P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Malatrasi,M., Miller,R.,
 Nguyen,H.T., Rausch,C.J., Seaton,C.L., Tong,J.C., and Zhang,D.

TITLE The structure and function of the expressed portion of the wheat
 genomes - Heat stressed spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773
Fax: 5105595818
Email: candersn@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: StrataGene SK primer.

FEATURES

Source

```

/organism="Trifolium aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1709_F10_119"
/clone_1lb="wheat heat stressed spike cdna library"
/issue_type="Whole spike"
/dev_stage="Spikes at 5, 10, 15 and 20 days after
anthesis"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid:
Site 1: EcoRI; Site 2: XhoI; Spikes at 5, 10, 15 and 20
days after anthesis were heat stressed under two
conditions at Texas Tech University (D. Zhang in RTM Nguyen
lab): (1) at 38 C for 4 hours and (2) 5 days of cyclic
treatment of 38 C for 4 hours. Total RNA and poly(A) RNA
were prepared, a cDNA library was made, and the cDNA
clones were in vivo excised to give pluscript phagemids
in the TJ Close lab (Choi, Close, Fenton, Malatrasi) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

```

BASE COUNT	80 a	114 c	104 g	71 t	1 others
ORIGIN					

```
alignment_scores:
  Quality: 82.50
  Ratio: 1.528
  Percent Simlarity: 60.674
  Length: 89
  Gaps: 3
  Percent Identity: 28.093
```

```
alignment_block:
US-09-528-682-1 x BF429191 ..
```

Align seg 1/1 to: BF429191 from: 1 to: 370

[illegible]

seq_name: gb_est2:BF422183

seq. documentation block:	614 bp	mRNA	linear	EST 28-NOV-2000
LOCUS	BF422183			
DEFINITION	FM1_12_G12.D1.A003 Floral-Induced	Meristem 1 (FM1)	Sorghum	
	proliferum cDNA, mRNA sequence.			

ACCESSION	BF422183	GI:11410172
VERSION	BF422183.1	
KEYWORDS	EST.	
SOURCE	Sorghum propinquum.	
ORGANISM	Sorghum propinquum	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 614)	Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M., Marsala, C. and Pratt, L.H.	An EST database from Sorghum: floral-induced meristems unpublished (2000)		Contact: Cordonnier-Pratt MM Department of Botany

Department Of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmp@atcuga.edu
Sequences have been trimmed to exclude POLYA⁺ vector and regions below Pired quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 583
POLYA-No.

FEATURES	Location/Qualifiers
source	1. .614

```

/organism="Scorhum propinquum"
/db_xref="taxon:132711"
/clone_lib="Floral-Induced Meristem 1 (FM1)"
/note="Organ: Floral-Induced Meristem; Vector:
pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
EcoRI; mature plants were placed in a growth chamber for
15 days with 16 hr darkness and 8 hr light (flowering is
induced by short-day conditions); 16 days after being
returned to the greenhouse under natural long days during
late April/early May, meristems were harvested. The
library was made from poly(A RNA in the cloning vector
lambda Zap II. Clones to be sequenced were prepared by
mass excision."
BASE COUNT      195 a      148 c      177 g      94 t

```

```
alignment_scores:      Length: 1111
                        Quality: 82.50
                        Ratio: 1.352    Gaps: 8
Percent Similarity: 54.955   Percent Identity: 32.432
```

```
alignment_block:
US-09-528-682-1 x BF422183 .
```

Align seg 1/1 to: BF422183 from: 1 to: 614

```

127  TRYR.....ArgValAspMethyleValIleAspGluArg 139
      |||||      ||      :::::
178  TGGTACCGACACAGCAGTAACTAGCTACTCTCCGCACTTTGACAAAGAGT 227
      |||||      |||||      |||||
139  uHisArg.....AsnArgGluTyrArgAspArgTyrT 150
      :::::      :::::      |||||
228  CCATGACAGGTATGTCGGAGGGTCCGACGACCGACGGCGAGGTC 277
      |||||      |||||      |||||
150  YrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGI 166
      |||||      |||||      |||||
278  GAGCGCCTCTACGAGGAGCGCCACAGAGGAGGACAGAGTCTCTCGGG 327
      |||||      |||||      |||||
166  yPheArg...ProAspHisGlnAlaTyrArgGluGluProTyrIleHis 181
      |||||      :::::      |||||
328  GAGACCCCGGACGAGGTGATGACGAGTGTGAAGTGCACCCCTGCAGCCAC 377
      |||||      |||||      |||||
182  HisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspTh 198
      |||||      |||||      |||||

```


[illegible]

seq. documentation block: 1159 bp mRNA linear EST 12-SEP-2001

LOCUS B1658957 603301833F1 NIH_CGAP_Mam4 Mus musculus CDNA clone IMAGE:5347497.5, mRNA sequence.

DEFINITION

ACCESSION B1658957

VERSION B1658957.1 GI:15573193

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1159)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LAM11883 row: b column: 10
High quality sequence stop: 178.

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="5347497"
/clone_id="NIH_CGAP_Man4"
/tissue_type="tumor, gross tissue"
/lab_host="DHI0B"
/notes="Organ: mammary; Vector: pCMV-SPOB6; Site:1: Nct1;
Site_2: Sal1; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lothar Hennigshausen/Priscilla Furch, NIH
Reference for transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NCL_CGAP
Library."
BASE COUNT      438 a      174 c      297 g      250 t
ORIGIN

```

```

alignment_block:
US-09-528-682-1 x B1658957/rev ..
Align seg 1/1 to reverse of: B1658957 from: 1 to: 1159

```

19 SerGlyGlyLeuMetProArgGlyHis..... 27
 ||| ||| ::||| |||
 774 TCCACGACCTCTTCCTCCACCATCCCTATCGTCACAGTCTTACCTG 725

```

28 ... AsnGluTyrPheAspArgIlyThrGlnMetAsnIleAsnIleuTYA 43
      :::::::::: :::::::::: ::::::::::
724 GCGGCATCTTACTTCGATCGATCTAATACATCTACAGATATCATCAC... 679
      :::::::::: :::::::::: ::::::::::
43 sPhisAlaArgIlyThrGlnThrGlyPheValArgTyrAspArgIlyTyr 59
      :::::::::: :::::::::: ::::::::::
678 ..CATATACGCGACATCTCTCCCTTCACGCGTTTCAC.....TAT 637
      :::::::::: :::::::::: ::::::::::
60 ValSerThrSerLeuSerLeuArgSer.....Ser 68
      :::::::::: :::::::::: ::::::::::
636 TTATCTACTAGCATAGACACTCACTTCATTCACACTCCACTCCATCAT 587
      :::::::::: :::::::::: ::::::::::
68 ..... 68
586 TACCTATATACGTACCTCATCTATGATGTATATGTCGCTCTCTCT 537
      :::::::::: :::::::::: ::::::::::
69 .....AlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThr 82
      :::::::::: :::::::::: ::::::::::
536 ACATACCTTCGACACTTATACCTCATCTTACTTCATACATCGGACTGTATC 487
      :::::::::: :::::::::: ::::::::::
83 TyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAs 99
      :::::::::: :::::::::: ::::::::::
486 TATCGCGCATATCTGTGCGCTATCTCTCGCTTCGCTTCACTATCCGTCAC 437
      :::::::::: :::::::::: ::::::::::
99 PAlleuGlyValTyr.....SerProHisProTyrGlnGlnGly 113
      :::::::::: :::::::::: ::::::::::
436 CATCTATATAGCTTTATACTTCTACATCTCCACAC.....CACC 399
      :::::::::: :::::::::: ::::::::::
113 AlSerAlaLeuGlyGlyIlePro...TyrSerGlnIleTyrGlyTyrTyr 128
      :::::::::: :::::::::: ::::::::::
398 CTTCATCTCTAGGACCTGCGCGCGCTCTGCATCTATATTTGTTACTACT 349
      :::::::::: :::::::::: ::::::::::
129 ArgValAsnPheGlyValIleAspGlnAlaGlyLeuHisArgAsnArgGly 145
      :::::::::: :::::::::: ::::::::::
348 .....GTCGCTCTCCATCTATATACATCACCTTACCTAGTA 314
      :::::::::: :::::::::: ::::::::::
145 fArgAspArgTyrTyrArg 151
      :::::::::: :::::::::: ::::::::::
313 CAATTATTCAGTACTTTCGC 295

```

seq_name:	gb_estl.BB618505
seq_documentation_block:	
LOCUS	BB618505 596 bp mRNA linear EST 31-AUG-2001
DEFINITION	BB618505 RIKEN full-length enriched, 8 days embryo Mus musculus
ACCESSION	CDNA clone 5730407E15 5', mRNA sequence.
VERSION	BB618505
KEYWORDS	BB618505.1 GI:15396603
SOURCE	EST.
ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 596)
AUTHORS	Arakawa,T., Carinini,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouoda, 'M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sasaki, Okazaki,Y., Okido,T., Salto,R., Sakai,C., Sakai,K., Sano,H., Sasaki, 'D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: <http://genome.qsc.riken.go.jp/>
carninci.1.p., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

M, Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagl, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matschik, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Konno, S., Saito, T., Shingawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Araiawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Func. Genomics* 2 pre, 172-186 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

Location/Qualifiers
1. .596

1. .596
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5730407E15"
/clone_1bp="RIKEN full-length enriched, 8 days embryo"
/sex="mixed"
/dev_stage="8 days embryo"
/lab_host="DH10b
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genetic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15',
GAGGAGAGAGAGATCTCAGAGAGCTCTTTTCTTTTCTTTT 3']. cDNA was prepared by using Trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-triapper. cDNA went through one round of subtraction to Rot = 100.0 Second strand cDNA was prepared with the primer adapter of sequence 15',
GAGGAGAGAGTCTCAGAGTAAATTAATTAATTAATCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUCS(+ after bulk excision from LambdaFLC I Cloning sites, 5' end: SalI, 3' end: BamHI."
175 c 127 g 126 t

82.00	length:	156
1.000	Gaps:	11
53.247	Percent Identity:	25.974

B618505

BB618505 from: 1 to: 596

...cugInglIvalSerAlblenclygly.... 118
::: ||| :::: :|||
::: ||| :::: :|||
CCCGCAGCTGCGCCAGACAAACTTAAGTCAGGAGCAGC 155
LeProTyISerGlnIleTyrGlyTrpTYrArgValA 131
||| ||||||::: |||
::: |||
CATATTCAGACGCTCCACAGCTAAGTGGAAGCACTCACCTTCGTA 205

131 snrhe1yua1lleasrgluatgleniaargvsnarvgluutytatgaar 147
||:::|||||
206 acsgaaat.....tcagacacagcagadagaactgtacgtaccagaaagt 249
148 argtutttatgaaaleuasnllea1aroadlaqluaspclutytargle 164
::: |||::: ||
250gagaaagtctgtctcggtgcgtgat...ttcagtcst 284
164 unlaglyphreprotopasp.....nsglnalatpragslug 177
| ||| |||||
285 gttggcgtcacccacscacttttcagacagcstactctcagatccaagac 334
177 luertotpllenishisala.....proglnglcysglyasn 189
::: ||| :::
335 agagttatttcacctccacagcaaaaatggccccgcgtacttgga... 381
190 serserargthrlierthy.....aspth 198
||| |||
382agcagaaacagcccaaagcttcttcacaccaggtggccacgtaccac 428
198 rcysaangluglutthrghlansleuserthrlietgylueargslutyc 215
|::: ||| :||::: |
429 tggcatgcacaattttttcagacaccttacacacattttggcttaggaattccac 478
215 inser.lysvallys...argglnilepheserasyptelnsrglua 230
||::: |||||
479 aaatatccgcatacagtgacacaaagctaacttcagattaacctcagggaagt 528
230 laspielietyr 233
|||::: |||
529 agatgtgttc 538

name: gb_gss:AZ247543

	AZ247543	613 bp	DNA linear	GSS 15-JUN-2000
DESCRIPTION	RPCI-23-92P2.TJ	Mus musculus genomic clone	RPCI-23-92P2,	
DEFINITION	DNA sequence.			

ION A2247543.1 GI:8560746
ORDS GSS.

SANITISM

REFERENCE AUTHORS

Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-92P2.TV

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.html>) or from Reseach Genetics (<http://www.resgen.com>). BAC end page: http://www.tigr.org/tdb/Bac_ends/mouse/bac_end_intro.html
Plate: 92 row: P column: 2
Seq primer: Sp6
Class: BAC ends.

JRES

```
1. .613
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
```



```

/clone="RPC1-23-92P2"
/clone_lib="RPC1-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC63.6; Site_1:
ECORI; Site_2: EORI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EORI and EORI Methylase. Site
selected DNA was cloned into the pBAC63.6 vector at the
EORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      172 a      131 c      136 g      174 t
ORIGIN

alignment_scores:
      Quality:      82.00      Length:      180
      Ratio:      0.943      Gaps:      12
      Percent Similarity: 48.333      Percent Identity: 24.444

alignment_block:
US-09-528-682-1 x AZ247543/rev ..

Align seg 1/1 to reverse of: AZ247543 from: 1 to: 613

41 LeuTyraSPHisAlaArgGlyThrGlnThrGlyPheValArgTyraSPas 57
      ::::| | | | | | | | | | | | | | | | | | | | | |
475 AACTATTTCATGTAGAAATCTACCAAGCAAGTGTATGATGAGGA 426
      ||| | | | | | | | | | | | | | | | | | | | | |
57 pGly.TyrValSerThrSerLeuArgSerIleHisLeuAlaGly 73
      ||| | | | | | | | | | | | | | | | | | | | | |
425 TGGATTATATC.....CTTGCTTGATGACGATTTCTATTTTGGG 385
      GlnSerIleLeu.....SerGlyTyrSerThr 82
      ||| | | | | | | | | | | | | | | | | | | | | |
384 CAACAATCATCATGATCAAGAAGCAAGTTGGGAGGAAGGTTATTACAG 335
      82 rTyTyTyTyTyTyValIleAlaThrAlaProAsnMetPheAsnValAsn 99
      T.....TACATTTCATTTGAGGAAGCAAGTCAAGACT..... 304
      99 spValLeuGlyValTyrSerProHisProTyrGlnGlnGluValSerAla 115
      ||| | | | | | | | | | | | | | | | | | | | | |
303 .....GGAAGTCAAGCAGCTCAGTCGAGGACGACGAGCTGATGACG 262
      116 LeuGlyGlyIleProTyrSerGlnIleTyrGlyTyTyTyTyValAsn 132
      ||| | | | | | | | | | | | | | | | | | | | | |
261 AGG.....CCATGGAAGGATGTTCAATTATTTGGCTT...GCCTCACC 224
      132 eGlyValIleAspGluArgLeuHisArgAsnArgGlyTyTyTyTyAspArg 149
      ||| | | | | | | | | | | | | | | | | | | | | |
223 CGGCTTGCTGACCTGCAGCTCTATGAAACCCAGACCTTCACCCAGA 175
      149 yTyTyTyTyArgAsnLeuAsnIleAlaProAlaGluAspGlyTyTyTyTyAla 165
      ||| | | | | | | | | | | | | | | | | | | | | |
174 ..AATAGTAACACCCCAAGGGGCT..... 151
      166 GlyPheProProAspHisGlnAlaTrpArgGluGluProTrpIle..... 180
      ||| | | | | | | | | | | | | | | | | | | | | |
150 .....TTTCCCTTGATCATCTAATTGAAATGCGCTTACAGCTGAGTCATANA 104
      181 .HisHisAlaProGln.....GlyCysGlyAsnSerSerArg 193
      ||| | | | | | | | | | | | | | | | | | | | | |
103 GAGGCACTTCCCACTGAAGCTCCTTCTCTGTGTATTAACCTTACG... 58
      193 hTrIleThrGlyAspThrCysAsnGluGluThrGlnAsn 205
      ||| | | | | | | | | | | | | | | | | | | | | |
57 .....CTGTGCAAGTTGACACCCAAAC 34
seq_name: gb_est1:BB619064
seq_documentation_block:
LOCUS      BB619064      634 bp      mRNA      linear      EST 31-AUG-2001

```

DEFINITION	BB6119064 RIKEN full-length enriched, 8 days embryo Mus musculus
ACCESSION	BB6119064
VERSION	BB6119064.1
KEYWORDS	GI:15396735
REFERENCE	EST.
AUTHORS	house mouse.
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
COMMENT	1 (bases 1 to 634) Atakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komou,H., Kouda,M., Koyu,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okito,T., Saito,R., Sakai,C., Sakai,R., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Atakawa,T., et al. 2001) Unpublished (2001) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komou,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagii,K., Fujiwaka,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Matsubara,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komou,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Yamataka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Atakawa,T., Ishii,Y. and Hayashizaki,Y. Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.
FEATURES	Location/Qualifiers
source	1..634 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="5730471122" /clone_id="RIKEN full-length enriched, 8 days embryo" /sex="mixed" /dev_stage="8 days embryo" /lab_host="DH10B" /note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGGAGAGAGAGATCCAGAGCCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by

BASE COUNT	182 a	170 c	147 g	135 t
ORIGIN				

alignment_block:

US-09-528-682-1 x BB619064

Align seg 1/1 to: BB619064 from: 1 to: 634

```

106 ProHsp70Ptyr.....GluGlnGluValSerAlaLeuGlyGly.....118
|||||.....:::|||||.....:::||||
53 CCACACCCACACCCGACACTGCCACAGACAAACTTAAGGTCCAGGACGAC 102
119 .....IleProTyrSerGlnIleTyrGlyTyrTyrGlyVala 131
||| |||||.....|||
103 ACTAGGTCACTTCATCTCCAGCTCCGACGTTAAGTGGCACTCAAGCTTCTGA 152
131 snPheGlyValIleAspGlyIleValGlyLeuIleValGlyAsnArgGlyTyrArgAsp 147
|||||.....:::|||||.....:::|||||
153 ACTGGAAAT.....TCAGGACAAGCAGAGAAAGAGTGTGAGCTCAAGAGAT 196
148 ArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgGlye 164
::: :::::|||||.....::: ||
197 .....GAGAACTGTGCTCTCTCTGGTGGCTGAT...TTCACTGT 231
164 uAlaGlyPheProProAsp.....HisGlnAlaIleTyrArgGlnG 177
| ||| |||||
232 GTTGGGCTTCACACCACTCTTCAGACAGCTCACTCTCCAGATCAAGAC 281
177 LuProTyrIleHisAla.....ProGlnGlyCysGlyAsn 189
::: ::: ||| ::::: |||||
282 AAGATTATTCACATCCACAGCAAGAAATGGCCCCGCTACTGTGGA... 328
190 SerSerArgThrIleThrGly.....AspThr 198
||||| |||
329 ...ACGAGAACAAAGCCAAAGCTTCTCTCTCCAGAGTGGCAAGCTGAC 375
198 rCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuValGlyTyr 215
|::: ::::: |||||
376 TCCCAATGCAAAATTTCATCGAGACCCATCAGCATTTGGCTCAGGAATCCAC 425
215 InSer LysValLys...ArgGlnIlePheSerAspTyrGlnSerLys 230
|::: ::::: |||||
426 AAAATTCGGATCAAGTGCACCAAGCTAACTTACGTACCTCAGGGAAGT 475
230 LAspIleTyr 233
|||||.....
476 ACATGTCTTC 485

seq_name: gb_gss:A2571452

seq_documentation_block:
LOCUS A2571452 687 bp DNA linear GSS 15-MAY-2001
DEFINITION 287Pv002 Pv MBN #30 Plasmodium vivax genomic 3', DNA sequence.
ACCESSION A2571452
VERSION A2571452.1 GI:13983551
KEYWORDS GSS.
SOURCE malaria parasite P. vivax.
ORGANISM Plasmodium vivax
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 687)

```

AUTHORS	Carlton, J.M. -R. and Dame, J.B.
TITLE	The Plasmodium vivax and P. berghel gene sequence tag project
JOURNAL	Parasitol. Today 16 (10), 409 (2000)
COMMENT	Contact: Dame JB

Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.

FEATURES	Location/Qualifiers
source	1. .687

```

/organism="Plasmodium vivax"
/srstrat="Salvador I (Collins, W. 1972. J. Parasitol. 69,
497-596)."
/db_xref="taxon:5855"
/clone_id="Py MBN #30"
/dev_stage="asexual blood forms"
/lab_host="Saimiri boliviensis"
/notes="Vector: pBluescript SK(+) vector DNA, phagemid
excised from lambda zap. Site_1: EcoR V; Site_2: EcoR V;
Host leukocytes were extracted from P. vivax infected
blood using the following methods: first, infected blood
was activated by the addition of 0.5 ml of ADP (40mg/ml)
per 10 ml blood. Then blood was passed over a column of
acid washed 0.1 mm glass beads, then through a Plasmidpur
filter, followed by passage through a column of pre-wet
Whatman CFI1 powder (1:2 ratio volume of blood to CFI1),
and finally centrifuged through a 50% Percoll density
cushion. Purified DNA was digested with mung bean nuclease
in the presence of 44% formamide at 500c, as described
(Vennick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). Digested DNA was
blunt-ended using T4 DNA polymerase and size fractionated
over a Sepharose CL-2B column. Fractions in the size range
500bp-4kb were ligated into the Eco RV site of pBluescript
SK(+), and E. coli XL-10 Gold transformed with the
ligation mixture."

```

BASE COUNT	178 a	169 c	168 g	170 t	2 others
ORIGIN					

alignment_scores:		
Quality:	82.00	length: 206
Ratio:	0.837	Gaps: 13
Percent Similarity:	47.573	Percent Identity: 24.757

```
alignment_block:
```

Align seg 1/1 to: A2571452 from: 1 to: 687

```

16  ILEYSAGSSEGLYLGLLeuMeProrgly.....Hisangl 29
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
16  TTTAAACCGAGCGGGGGGTACACCCCGAGAGACTTATCAGAAAAAAA 65
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
29  uTyrPheaspArgGLYThrGlnMetAsnIleasnLeuTyrAspHisAla 46
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
66  AACATTTCAAACGGGGCGCC.....ATTCAAACGAGCGGCATTTCAA 108
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
46  rgGLYThr.....GlnThrGlyPheValArgTyr 55
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
107  ACGGGGGCGTCATTCAATATGGGGCGCCCCCAAGAGAAC.....CGGTAT 150
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
56  AspAsp.....GlyTyrValSerThrSerLeuSerLeu.. 66
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
151  CCAACGAAACGGCCATCCAAACGAACGCCCGCGTCAGCGCGGCTGGG 200
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
67  .....ArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerIlyr 80
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|
201  GAGACACATGCTTTCGCGCCCTGGAGACGAAGACTCTGCTGTATTGGGG 250

```



```

80  yrSerThrTyrTlleTyrValIleAlaThrAlaProAsnMetPheasn 96
    |||::: |||:::
251 GTGGCAGCTATCTCAAGCAATCCAGCTTTGCCCTTATTCAC 300
97  ValAsn.....AspVa 100
    ::|||
301 CTCACGCTTAACTGAACCCGACTTGAAGTCCGCTTTGCACGAGAGT 350
100  lleuGlyValTyrSerPro.....HisPro.....T 109
    |::: ||| ||| ||| |||
351 AGTTGGTACGCTGCCCGTTGTCATGAAAGCTATCCAGTGTCT 400
109  yrgIuGlu.....ValSerAlaLeu 116
    |||::: |||
401 ATGAGGAAGAACCGCCATTCAGTAGAGCTTATTCATGAGTGTG 450
117  GlyGlyIleProTyrSerGlnIleTyrGlyTyrTyrAlaValAsnPheG1 133
    ::|||::: |||::: |||:::
451 TCTGCTGTCTCCACGACCAATG.....CTTNG 479
133  yValIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrT 150
    ::: ::: ||| ||| |||
480 CCTTNGTAGTGTGTAATAATTAGAAATATGGAACAGGCTTTTATATC 529
150  yratGAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeu.AlaG1 166
    ::::: |||::: ||| ||| |||
530 ATAAGGAGCTCTCCTTGAGA.....TATGCTCTCAAGAG 564
166  yPheProProAspHis 171
    ::||| ||| ||| |||
565 ACTCCCTCCGACCTAC 580
seq_name: gb_est2:BG539191

```

```

seq_documentation_block:
LOCUS      BG539191          832 bp      mRNA      linear      EST 03-APR-2001
DEFINITION 602566883F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691349 5',
            mRNA sequence.
ACCESSION  BG539191
VERSION     BG539191.1 GI:13531424
KEYWORDS   EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 832)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1510 row: f column: 22
            High quality sequence stop: 495.
            Location/Qualifiers
                1..832
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4691349"
                /clone_lib="NIH_MGC_77"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site:1:
                5' adaptor used in cloning as follows: 5' adaptor
                sequence: 5'-CACGGCATTATGAGCC-3' and 3' adaptor sequence:
                5'-ATTCTAGAGCGCGAGCGCGGCACATG-3' (where B = A,

```

C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

```

BASE COUNT      212 a      240 c      223 g      157 t
ORIGIN
alignment_scores:
    Quality:      82.00      Length:      103
    Ratio:        1.491      Gaps:      7
    Percent Similarity: 53.398      Percent Identity: 33.010
alignment block:
US-09-528-682-1 x BG539191

```

Align seq 1/1 to: BG539191 from: 1 to: 832

```

111 GlnGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyT 127
    |||::: |||::: |||::: |||::: |||::: |||::: |||:::
489 CAGCGGTGTCAGCGGCTCCGCGCTG...TGCTCCAGGCATATGAGAG 535
127 pTyrArgValAsnPheG1ValIleAspGluArgLeuHisArgAsnArg 144
    |||::: |||::: |||::: |||::: |||::: |||::: |||:::
536 GCCGAGGCTGGGC.....ACGATCCACAGAGTCTGGG 567
144 lutyArgAspArg..... 148
    |||::: |||::: |||::: |||::: |||::: |||::: |||:::
568 AGTACAAGACCGCTGCGACATGAGTGAACCCGCTTCTACCTAA 617
149 ...TyrTyrArgAsnLeu.....AsnIleAlaProAlaGluAs 160
    ::: |||::: |||::: |||::: |||::: |||::: |||:::
618 ACACACTAGCGGGGCTGTGTGACACCTGTAGTGGCGGACAGACG 667
160 pGlyTyrArgLeuAlaGlyPhePro...ProAspHisGlnAlaTTrArg 176
    |||::: |||::: |||::: |||::: |||::: |||::: |||:::
668 GGGGGCATGCGCGCAACCCGCTGACCCGAGGAGAGCTGACGACAAAG 717
176 lugiProTPrIleHisHisAlaProGlnGlyCysGly.....AsnSer 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
718 CGGAG.....CATCACAGGCCCGCGGCGGCGCAAAAGATTC 758
191 SerArgThr 193
    ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 GCCCGCAC 767
seq_name: gb_est2:BF304330

```

```

seq_documentation_block:
LOCUS      BF304330          1935 bp      mRNA      linear      EST 21-NOV-2000
DEFINITION 601887226F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4121047 5',
            mRNA sequence.
ACCESSION  BF304330
VERSION     BF304330.1 GI:11251055
KEYWORDS   EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1935)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: L1CM1002 row: d column: 08

```


High quality sequence stop: 19.
location/Qualifiers
1..1935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4121047"
/clone_lib="NIF.MGC.17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOT57; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoR/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 526 a 573 c 471 g 360 t 5 others

ORIGIN

alignment_scores:
Quality: 82.00 Length: 92
Ratio: 1.577 Gaps: 5
Percent Similarity: 56.522 Percent Identity: 29.348

alignment_block:
US-09-528-682-1 x BF304330 ..

Align seg 1/1 to: BF304330 from: 1 to: 1935

6 TTTATG.....AlaAspSerArgProProAspGluIleLysArgSerGI 20
|||||
|||||
1286 TACCGCATACCCGCTGACGCGCAGTTCGCGCTACTGATCAGCCCTTATGG 1335
20 YGILeu.....MetProArg 26
|||||
1336 TGGCTTACTTTTCGCCCTACCTCATATGTTAGTACCACCACATACCTCGC. 1384
26 LYLHLSAnGIuTTrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyr 42
|||||
1385 ..CATGCGAATTGGATGAGACAAAGAAAGAACCCCTTAGTATA.....CtC 1426
43 AsPHisAlaIArgGlyThrGlnThrGlyPheValArgTyrAspArgLty 59
|||||
1427 GACCAACACACGCTAGACAGACAAACGGTTTCACAAACACCCCATGGAAA 1476
59 rValSerThrSerLeuSer.....LeuArgSerAlaHisLeuAlaGlyc 74
|||||
1477 CATGCGCGACAAACGGAGTGACATTAACAATCTCAACACACACGACAGCA 1526
74 InSerIleLeuSerGlyTyrSerThr 82
|||
1527 CGGCATGTGCGACGACGACACAGAAC 1552
seq_name: gb_est2:BE864213

seq_documentation_block:
LOCUS BE864213 368 bp mRNA linear EST 29-SEP-2000
DEFINITION UT-M-BH1-ant-c-02-0-UT_r1 NIH_BMAP_M_S2 Mus musculus cDNA clone
UT-M-BH1-ant-c-02-0-UT 5', mRNA sequence.
VERSION BE864213
KEYWORDS BE864213.1 GI:10385033
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 368)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

```

MEDLINE
COMMENT      97044477
              Contact: Chin, H
              National Institute of Mental Health
              6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
              20892-9643, USA
              Tel: 301 443 1706
              Fax: 301 443 9890
              Email: MEST@mail.nih.gov
              cDNA library preparation: M.B. Soares Lab Clone distribution:
              Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
              should be noted that Bento Soares is generating a small number of
              additional specialized non-redundant arrays of BMAP cDNAs whose
              availability will be considered under appropriate and limited
              collaborative arrangements
              Seq primer: M13 Reverse.
FEATURES
SOURCE       location/Qualifiers
              1. 368
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="U1-M-BM1-ant-c-02-0-U1"
                /clone_11d="NIH_BMAP_M_S2"
                /dev_stage="27-32 days"
                /lab_host="DH10B (Life Technologies)"
                /note="Vector: pR773D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; The
                NIH_BMAP_M_S2 library is a subcloned library derived from
                NIH_BMAP_M_S1, which in turn is a subcloned library
                derived from a mixture of normalized libraries from ten
                regions of the mouse brain (cerebellum, brain stems,
                olfactory bulbs, hypothalamus, cortex, amygdala, basal
                ganglia, pineal gland, striatum, hippocampus). The driver
                used for subtraction consisted of a pool of 5,000 clones
                from the NIH_BMAP_M_S1 library and a pool of 2,000 clones
                obtained from non-normalized and normalized mouse brain
                spinal cord libraries."
BASE COUNT   83 a 98 c 94 g 93 t
ORIGIN
Alignment_scores:
              Quality: 81.50          Length: 132
              Ratio: 1.273           Gaps: 8
              Percent Similarity: 48.485   Percent Identity: 25.758
alignment_block:
US-09-528-682-1 x BE864213 ..
Align seg 1/1 to: BE864213 from: 1 to: 368
101 LeuGlyValTyrSerProHisProTyrGlu.....GlnGluValSe 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
10 CTTGTGCTGTTCTCCCAACACCCCAATCCAGAAATCTTCACAGATCTCA 59
114 ValAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly..TrrTyr..A 129
    ::|||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 CAGTGTGATGTGTTACCCCTACATGTCATCATGATGAGACGTGTTCTGNG 109
129 rGValAsnPhcGlyValIleAspGluArgLeu..... 139
    ::| |||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
110 GGAAGCTGTGGGGCGTCTCCCGCTCAAGTGGAGCTGTGCTCAAT 159
140 .....HisArgAsnArgGluTyr.....ArgAspArgTyr 149
    ||| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 GCCTACGTGATCATCTACATGCTGAGTACACCCAGACAGAAAGACATCTCA 209
149 rTyrArGAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaG 166
    ||::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 CTTT.....GCACACGCTGTGGCC. 228
166 LYPheProAspAspHisGlnAlaATrPAArgGluGluProTrpIleHisHis 182
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 .....CAAGCTTGGGAAGTGGCCCGACAGTTCATCCACACGC 261

```



```

183 AlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 199
    ::::: |||::: :::::
262 ACATCCAAAC.....AATGCAGATGCGTTCATATTGTGCGAGACCT 302
    |||::: |||::: |||:::
199 sAsnGluGlnThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
    |||::: |||::: |||:::
303 CAATATGCAACCCCAAGACCTGGCTGCTGCTGCTGCAAGAGTGG 348

seq_name: gb_est2:BJ180858

seq_documentation_block:
LOCUS      BJ180858          489 bp      mRNA      linear      EST 24-JAN-2002
DEFINITION BJ180858 normalized full length cDNA library, chloronemata,
            caulonemata and malformed buds Physcomitrella patens subsp. patens
            cDNA clone pphb29a23 5', mRNA sequence.
ACCESSION   BJ180858
VERSION     BJ180858.1  GI:18348809
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens.
ORGANISM    Physcomitrella patens subsp. patens
REFERENCE   Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
AUTHORS     Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
            1 (bases 1 to 489)
            Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,
            Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
            ,M.
            Comparison of the moss Physcomitrella patens genome with flowering
            plants genome
            Unpublished (2002)
            Contact: Tadasu Shin-I
            Center For Genetic Resource Information
            National Institute of Genetics
            111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp
            A backbone of the vector is Bluescript II, that was in vivo
            excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI
            digested-5' end of cDNA is ligated to SalI site of the vector, and
            the BamHI digested-3' end including poly-A tail is ligated to BamHI
            site of the vector. cDNA insert could be amplified with
            conventional T7 and T3 primers. This normalized full-length cDNA
            library was generated basically according to the method described
            in Genome Research 10, 1617-1630 (2000), Carninci, P. et al.
            Protoneemata were blended by the POLYTRON, and then cultivated on
            the BCD medium containing 0.5um BA (benzylaminopurine) for 8 to 13
            days under the continuous light.
            Location/Qualifiers
                1. 489
                /organism="Physcomitrella patens subsp. patens"
                /db_xref="taxon:145481"
                /clone="pphb29a23"
                /clone_lib="normalized full length cDNA library,
                chloronemata, caulonemata and malformed buds"
                /tissue_type="mixture of chloronemata, caulonemata and
                malformed buds"

BASE COUNT      115 a      112 c      145 g      117 t
ORIGIN
alignment_scores:
    Quality:      81.50      Length:      70
    Ratio:         1.772      Gaps:         4
    Percent Similarity: 65.714      Percent Identity: 34.286

alignment_block:
US-09-528-682-1 x BJ180858/rev ..

Align seg 1/1 to reverse of: BJ180858 from: 1 to: 489
|||||::: ::|||:::|||||
85 ||eTyrValIleAlaThrAlaProAsn.....MetPheAsnYA 97
|||||::: ::|||:::|||||

```

```

296 ATCTTCACAGAAACCACTCCGCCGAACATGCATGAAGAATCTTTTACCT 247
    |||::: |||::: |||::: |||:::
97 lAsnAspValLeuGlyValTyrSerProHisProTyrGluGlnGluValS 114
    |||::: |||::: |||::: |||:::
246 CTCTTCGGGCTTCATATGACAGCCCAACGACGATCAGTGAATGGTGT 197
    |||::: |||::: |||::: |||:::
114 eAlaLeu.....GlyGlyIleProTyrSerGlnIleTyrGlyTyrPyr 128
    |||::: |||::: |||::: |||:::
196 CGTCATCTCACTGGGAATGAATGAACCTCTTCCGAG.....TTTAT 156
    |||::: |||::: |||::: |||:::
129 ArgValAsnPheGlyVal.IleAspGluArgLeuHisArgAsnArgGluT 145
    |||::: |||::: |||::: |||:::
155 CCGTCATATTCGCGATAGTAGAGCAATCATCTCATCAGAGAAATCTC 106
    |||::: |||::: |||::: |||:::
145 YrArgAsp 147
    |||::: |||::: |||::: |||:::
105 TCGAGGAT 98

seq_name: gb_est2:BJ187033

seq_documentation_block:
LOCUS      BJ187033          608 bp      mRNA      linear      EST 24-JAN-2002
DEFINITION BJ187033 normalized full length cDNA library, chloronemata,
            caulonemata and malformed buds Physcomitrella patens subsp. patens
            cDNA clone pphb39116 5', mRNA sequence.
ACCESSION   BJ187033
VERSION     BJ187033.1  GI:18354974
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens.
ORGANISM    Physcomitrella patens subsp. patens
REFERENCE   Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
AUTHORS     Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
            1 (bases 1 to 608)
            Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,
            Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
            ,M.
            Comparison of the moss Physcomitrella patens genome with flowering
            plants genome
            Unpublished (2002)
            Contact: Tadasu Shin-I
            Center For Genetic Resource Information
            National Institute of Genetics
            111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp
            A backbone of the vector is Bluescript II, that was in vivo
            excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI
            digested-5' end of cDNA is ligated to SalI site of the vector, and
            the BamHI digested-3' end including poly-A tail is ligated to BamHI
            site of the vector. cDNA insert could be amplified with
            conventional T7 and T3 primers. This normalized full-length cDNA
            library was generated basically according to the method described
            in Genome Research 10, 1617-1630 (2000), Carninci, P. et al.
            Protoneemata were blended by the POLYTRON, and then cultivated on
            the BCD medium containing 0.5um BA (benzylaminopurine) for 8 to 13
            days under the continuous light.
            Location/Qualifiers
                1. 608
                /organism="Physcomitrella patens subsp. patens"
                /db_xref="taxon:145481"
                /clone="pphb39116"
                /clone_lib="normalized full length cDNA library,
                chloronemata, caulonemata and malformed buds"
                /tissue_type="mixture of chloronemata, caulonemata and
                malformed buds"

BASE COUNT      156 a      158 c      169 g      125 t
ORIGIN
alignment_scores:
    Quality:      81.50      Length:      70
    Ratio:         1.772      Gaps:         4

```



```

DEFINITION 602759363F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4895030 5',
RNA sequence.
ACCESSION B1199325
VERSION B1199325.1 GI:14654346
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 807)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cga@rs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
            Plate: L10M1779 row: m column: 15
            High quality sequence stop: 798.
FEATURES
    source
        1..807
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4895030"
            /clone_lib="NIH_MGC_19"
            /tissue_type="neuroblastoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Library constructed by Ling Hong
            in the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC library."
BASE COUNT 182 a 232 c 213 g 180 t
ORIGIN
alignment_scores:
    Quality: 81.00 Length: 154
    Ratio: 1.095 Gaps: 8
Percent Similarity: 48.052 Percent Identity: 25.974
alignment_block:
US-09-528-682-1 x B1199325/rev ..
Align seg 1/1 to reverse of: B1199325 from: 1 to: 807
61 SerThSerLeuSerLeuArgSerAlaHisLeuAlaGlyInSerIleLe 77
   :::::::::::::::::::: ::::::::::::::::::::
487 GCACACGAGATGTCAGTGCATTCACATCTTCGTCAGGAAT..... 443
77 userGlyTyrSer.....ThrTyrTyrI 85
   ::::::::::::::::::::
442 .TCTGGGTACGGACACGAGAGTCAAAATCTGTGGGATTTTCTCATTTGCC 394
85 leTyrValIleAlaThrAlaProAsnMetPheAsnValAsnaspValLeu 101
   ::::::::::::::::::::
393 CTATATCTCCGTCGTCACCATCT..TATAGTGTCCACAGAGTGTG 347
102 GlyValTyrSerProHisProTyrGluGlnGlnValSerAlaLeuGly.. 117
   ::::::::::::::::::::
346 GGGCCGACGAGATCTCATCCACGACGCGCAGCTGTACCATGTGCTC 297
118 .....GlyIleProTyrSerGlnI 124
296 ATCTCAACACCCACCCACCATGATTTCATTTGGGGAAGCCATTGATC.... 251

```

```

124 leTyrGlyTyrTyrAlaValAsnPheGlyValIleAspGluArgLeuHis 140
   ::::::::::::::::::::
250 .....TTCAGTACTGCTCAGAGTAGTCTGAGACTCCTCCG 212
141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaPr 157
   ::::::::::::::::::::
211 AAGTACTGGGGCTAC...GGAGAGCTGTATCCAAACTGTCATACCAAC 165
157 oHa.....GluAspGlyTyrA 163
   ::::::::::::::::::::
164 GGCACATCGGGGGTCCCGGGGGTCAACACATACAGATGTGTGTCATT 115
163 rGleuAlaGlyPheProProAspHisGlnAlaThrParGluGlnProTyr 179
   ::::::::::::::::::::
114 TTCTGGCAGAGGTCCA.....CATGTCGACAGAACAGGCAATCCCACT 71
180 lIeHisHisAla 183
   ::::::::::::::::::::
70 CTCATCACGCA 59
seq_name: gb_est2:B1152296
seq_documentation_block:
LOCUS B1152296 1401 bp mRNA linear EST 05-JUL-2001
DEFINITION 602917727F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5068039 5',
RNA sequence.
ACCESSION B1152296
VERSION B1152296.1 GI:14612297
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1401)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cga@rs-remail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: L10M1183 row: f column: 08
            High quality sequence stop: 153.
FEATURES
    source
        1..1401
            /organism="Mus musculus"
            /strain="C57BL/6J (f1)"
            /db_xref="taxon:10090"
            /clone="IMAGE:5068039"
            /clone_lib="NCI_CGAP_Lu29"
            /tissue_type="spontaneous tumor, metastatic to mammary.
            stem cell origin."
            /lab_host="DH10B"
            /note="Organ: Lung; Vector: pCMV-Sport6; Site_1: SalI;
            Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
            Library constructed by Life Technologies. Investigator
            providing samples: Gilbert Smith, NIH"
BASE COUNT 466 a 568 c 246 g 121 t
ORIGIN
alignment_scores:
    Quality: 81.00 Length: 208
    Ratio: 0.818 Gaps: 11
Percent Similarity: 47.596 Percent Identity: 24.038
alignment_block:
US-09-528-682-1 x B1152296 ..

```



```

Email: cgapds-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.jiml.gov
plate: L1AM9424 row: j column: 11
High quality sequence start: 3
High quality sequence stop: 299.
Location/Qualifiers
1. 1576
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4154602"
/clone_id="NCI_CGAP_Brn67"
/lisue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (rt phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NCI;
Site_2: Salt; Cloned unidirectionally. Primer: oligo dT
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 592 a 472 c 375 g 137 t
ORIGIN
alignment_scores:
Quality: 81.00 Length: 150
Ratio: 1.052 Gaps: 5
Percent similarity: 51.333 Percent identity: 22.000
alignment block:
US-09-528-682-1 x BF345560 ..
Align seg 1/1 to: BF345560 from: 1 to: 1576
79 gIyTyrSerThrTyrTleTyrlleTyValIleAlaThrAlaPro..... 92
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
138 GGAATCGATTCGTATGCTGCTGTCATTCCTAGAACGGCATCCACTTGTCTC 187
93 ..... AsmethepheaInValnsapValL 101
188 AGACCCCATTTGCTGTATGCATGACAGTAAGACGGTGTATGCATTCATCCATGC 237
101 euGIValIyTyrSerProHISProTyrgInglInValSerAlaLeuIly 117
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
238 TCGGCACATGAGGATTATATCCATTAAGAATCAACAGAGAACGGCAG 287
118 gIyIleProTyrSerGlnIleTyrgIyTrrPyTArgValaInPheGlyVa 134
||||| : : : : : ||| : : : : : : : : : : : : : : : : :
288 GAAGACCCCTTACTACATGACACAAAGG.....ATCAGATGAAGAT 328
134 IILeAsP.....GluArgLeuHISArgAsnArgGluTyrA 146
: : : : : : : : : : : : : : : : : : : : : : : : :
329 TACACATACACAGAAAGAACTAGATAGAAAGCCGTACACAAAGTACG 378
146 rGAsPArgTyrTyrArgAsnLeuAsnIleAlaProAlGIAsnAspGlyTyr 162
379 CCTG:CAATGGCAGCAAAATAGAACATCCGGAACATATGGAAGACGTACTAC 427
163 ArgLeuAlaGlyPheProProAspHisGlnAlaIleTrrParGluGluProTr 179
||||| ||| |||||
428 AGACTACAGGGGTGACCCACCG..... 448
179 pIleHISHisAlaProGlnGlyCysGlyAsnSerSerArgThrIleTyr 196
449 .....CACCGACCATATGCC.....AGCCCCACAGCAGCAATAG 482
196 IyAsPThrCysAsnGluIuThrGlnAsnLeuSerThrIleTyrLeuArg 212
|| : : : : : ||||| : : : : : ||| : : : : :
483 GACCCGCGCAAGGACGACAGACAGAAACAGACACTGGTACTATAGGA 532

```



```

seq_name: gb_est2:BF520534
seq_documentation_block:
LOCUS      BF520534                464 bp    mRNA    linear    EST 08-DEC-2000
DEFINITION  ESF458006 D51L Medicago truncatula cDNA clone pDSIL-24M2, mRNA
sequence.
ACCESSION   BF520534
VERSION     BF520534.1  GI:11609217
KEYWORDS    EST.
SOURCE      barrel medic.
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
            Medicago.
REFERENCE   1 (bases 1 to 464)
AUTHORS    Fedorova,M., Pleeson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
            ,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
            Holt,I.E. and Fraser,C.M.
            ESTs from leaves of Medicago truncatula after inoculation with
            Colletotrichum trifolii
JOURNAL     Unpublished (2000)
COMMENT     Contact: Deborah A. Samac
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
            Tel: 612 625 1243
            Fax: 651 649 5038
            Email: debbys@puccini.crl.umn.edu
            University of Minnesota name: M277975e
            TIGR sequence name: MTECV73JK
            More information is available at: http://chrysis.tamu.edu/medicago
            Seq primer: SKmod (CTA GAA CTA gta gta CC).
FEATURES
    source
        1..464
            /organism="Medicago truncatula"
            /cultivar="genotype A17"
            /db_xref="taxon:3880"
            /clone="pDSIL-24M2"
            /clone_lib="D51L"
            /tissue_type="Leaves infected with Colletotrichum
            trifolii"
            /dev_stage="cotyledons and primary leaves harvested 5 and
            8 days after inoculation with Colletotrichum trifolii"
            /lab_host="E. coli strain XLOLR"
            /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI; cDNA was prepared from polyA+ enriched RNA from
            cotyledons and primary leaves harvested 5 and 8 days after
            inoculation with Colletotrichum trifolii. The cDNA was
            directionally ligated into the Uni-ZAP XR vector from
            StrataGene and packaged using GigaPack III Gold packaging
            extracts. Plasmids containing cDNA inserts were excised
            from the recombinant lambda-ZAP phage using Ex Assist
            helper phage and propagated in XLOLR cells. Note: EST may
            be of fungal origin."
BASE COUNT      124 a      101 c      88 g      151 t
ORIGIN
alignment_scores:
    Quality:      80.50      Length:      72
    Ratio:        1.750      Gaps:      3
Percent Similarity: 63.889      Percent Identity: 33.333
alignment_block:
US-09-528-682-1 x BF520534/rev ..
Align seg 1/1 to reverse of: BF520534 from: 1 to: 464
2 GYAsPARGLeuTYrARGAlaSPSerArgProProlaspGluLeuYSar 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 GGACAAAGTTCTATGTAGTC.....CCATAAGACACATACGAAA 372

```

```

18 gserGlyLeuMeProlArgGlyHisAsnGluTYrPheAspArgLYT 35
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
371 GCTTCAGAGTCTTCTCCAGGAGACTCTCT.....TCTCCTGCTC 331
35 hrcGlnMetAsnIle...AsnLeuTYrAspHisAlaArgLYThrcGlnThr 50
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 GTGAATGACAAATCGGTCTTCTTACGATGAGACGTACACACAAATCC 281
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 CGCAGCATGAGCAAGACATGATGTATGATCTACTCTGCTATCTTAA 231
67 gserAlaHisLeuAla 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 GAAGAAACATATTGCG 215
seq_name: gb_est1:AU209914
seq_documentation_block:
LOCUS      AU209914                514 bp    mRNA    linear    EST 17-JUL-2001
DEFINITION  AU209914 unpublished oligo-capped cDNA library, stage I1
Caenorhabditis elegans cDNA clone YK750d06 3', mRNA sequence.
ACCESSION   AU209914
VERSION     AU209914.1  GI:14845824
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
ORGANISM    Caenorhabditis elegans.
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdicoidea
            ; Rhabdilitidae; Pelodierinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 514)
AUTHORS    Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
            A complementary view of the C.elegans genome
JOURNAL     Unpublished (2001)
COMMENT     Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykoha@lab.nig.ac.jp.
FEATURES
    source
        1..514
            /organism="Caenorhabditis elegans"
            /strain="N2"
            /db_xref="taxon:6239"
            /clone="YK750d06"
            /clone_lib="unpublished oligo-capped cDNA library, stage
            I1"
            /sex="Hermaphrodite"
            /tissue_type="whole animal"
            /dev_stage="I1"
BASE COUNT      97 a      118 c      154 g      145 t
ORIGIN
alignment_scores:
    Quality:      80.50      Length:      137
    Ratio:        1.184      Gaps:      7
Percent Similarity: 49.635      Percent Identity: 24.088
alignment_block:
US-09-528-682-1 x AU209914/rev ..
Align seg 1/1 to reverse of: AU209914 from: 1 to: 514
102 GlyValTYrSerProHisProTYrGlnGlnGluValSerAlaLeuGlyGcl 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
390 GGGATTCATCATCCGTCATCATCATCAACAACAA.....GA 356
118 yIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhcGlyValI 135
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


LOCUS	AU222397	533 bp	mRNA	linear	EST 17-JUL-2001
DEFINITION	AU222397 unpublished oligo-capped cDNA library, stage L1				
ACCESSION	Caenorhabditis elegans cDNA clone yk1015g07 3', mRNA sequence.				
VERSION	AU222397				
KEYWORDS	AU222397.1 GI:14860554				
SOURCE	EST.				
ORGANISM	Caenorhabditis elegans.				
REFERENCE	Caenorhabditis elegans.				
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;				
TITLE	Rhabditidae; Pelodertinae; Caenorhabditis.				
JOURNAL	1 (bases 1 to 533)				
COMMENT	Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.				
	A complementary view of the C.elegans genome				
	Unpublished (2001)				
	Contact: Yuji Kohara				
	Genome Biology Lab.				
	National Institute of Genetics				
	Yata 1111, Mishima, Shizuoka 411, Japan				
	Tel: 81-559-81-6854				
	Fax: 81-559-81-6855				
	Email: ykohara@lab.nig.ac.jp.				
FEATURES	Location/Qualifiers				
source	1..533				
	/organism="Caenorhabditis elegans"				
	/strain="N2"				
	/db_xref="taxon:6239"				
	/clone="yk1015g07"				
	/clone_lib="unpublished oligo-capped cDNA library, stage				
	L1"				
	/sex="Hermaphrodite"				
	/tissue_type="whole animal"				
	/dev_stage="L1"				
BASE COUNT	101 a	123 c	156 g	153 t	
ORIGIN					
alignment_scores:	Quality:	80.50	Length:	137	
	Ratio:	1.184	Gaps:	7	

Align seg 1/1 to reverse of: AU222397 from: 1 to: 533

78 ACCACAGCCGA 68

/db_xref="taxon:6239"


```

/cclone="yk749g06"
/cclone_lib="unpublished oligo-capped cDNA library, stage
L1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

```

BASE COUNT 106 a 130 c 163 g 159 t 1 others

ORIGIN

alignment_scores:

Quality:	80.50	Length:	137
Ratio:	1.184	Gaps:	7
Percent Similarity:	49.635	Percent Identity:	24.088

alignment_block:

US-09-528-682-1 x AU209839/rev ..

Align seg 1/1 to reverse of: AU209839 from: 1 to: 559

```

102 GlyValTyrSerProHisProTyrGluGluValSerAlaLeuGly1 118
      :::::|||||::: :::::|||||:::
394 GGATTCATCATCACCGACCATCATCAACAACAA.....GA 360
118 ylleProTyrSerGlnIleTyrGlyTrrPrrYrrarGvalAsnPhgIleVal1 135
      :::::|||||::: :::::|||||:::
359 TCACCGTATTGGAAAGTCTGCTCT.....CACTGAGGAAAGGA 322
135 leAspIuArGleuHisArGAsnArGluTyrArGAspArGTrYrrYrrArG 151
      :::::|||||::: |||:::|||||:::
321 AGAACAACAGCATCCACCGAGTTCGATCTTACCCAGAACACCATCACCC 272
152 AsnLeuAsnIleAlaProAlaGluAspGlyTrrArGleuAlaGlyPheP 168
      ||| :::::|||||::: |||:::|||||:::
271 AATGGAGAGATGCCAAGATACGG.....TATCGTCAACACGAGCTA.. 230
168 oProAspHisGlnAlaTrp.....ArgGluGluProTrpIleHis 182
      ||| |||:::|||||::: |||:::|||||:::
229 ....CAATGATGCTCGTGGAGCCGTTCTGGACCAAGACGCTTATC 184
182 lAsAlaProGlnGlyCysGlyAsnSerSerArGThrIleThrGlyAspThr 198
      || |||||::: :::::
183 ACCCTCCGCAAGTCGCTCATCACCCA..... 158
199 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArGluTrrG1 215
      |||:::|||||::: |||:::|||||:::
157 ...AACCAAGAGAGTCGCCCA.....CGAAGATATCA 129
215 nSerLysValLysArGlnIlePheSerAspTrrGlnSerGlnValAsp1 232
      | :::::|||||::: |||:::|||||:::
128 ACCTCAAGAGGATCGACACCATCTCCAAAGACGAGACGTCGCTTCAA 79
232 leTrrAsnArG 235
      :::::|||||
78 ACCACAGCCGA 68

```

seq_name: gb_est1:AU215136

seq_documentation_block:

LOCUS	AU215136	571 bp	mRNA	linear	EST 17-JUL-2001
DEFINITION	AU215136 unpublished oligo-capped cDNA library, stage L2				
ACCESSION	Caenorhabditis elegans cDNA clone yK821909 3', mRNA sequence.				
VERSION	AU215136				
KEYWORDS	AU215136.1 GI:14853293				
SOURCE	EST.				
ORGANISM	Caenorhabditis elegans.				
REFERENCE	Caenorhabditis elegans				
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidea				
	; Rhabditidae; Pelodierinae; Caenorhabditis.				
	1 (bases 1 to 571)				
	Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.				
	and Sugano,S.				
	A complementary view of the C.elegans genome				

[illegible]


```

DEFINITION      AU211806 unpublished oligo-capped cDNA library, stage L2
ACCESSION       AU211806
VERSION         AU211806.1 GI:14849353
KEYWORDS        EST.
SOURCE          Caenorhabditis elegans.
ORGANISM        Caenorhabditis elegans.
REFERENCE       Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Rhabditoidea
AUTHORS         ; Rhabditidae; Pelodierinae; Caenorhabditis.
                Kohara,Y., Shin-I,T., Thiery-Mieg,J., Thiery-Mieg,D., Suzuki,Y.
                and Sugano,S.
TITLE           A complementary view of the C.elegans genome
JOURNAL         Unpublished (2001)
COMMENT         Contact: Yuji Kohara
                Genome Biology Lab.
                National Institute of Genetics
                Yata 111, Mishima, Shizuoka 411, Japan
                Tel: 81-559-81-6854
                Fax: 81-559-81-6855
                Email: ykohara@lab.nig.ac.jp.
FEATURES
source          1..586
                /organism="Caenorhabditis elegans"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="yk776h04"
                /clone_1lb="unpublished oligo-capped cDNA library, stage
L2"
                /sex="Hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L2"
BASE COUNT     113 a 137 c 169 g 167 t
ORIGIN
alignment_scores:
                Quality: 80.50 Length: 137
                Ratio: 1.184 Gaps: 7
Percent Stmilarity: 49.635 Percent Identity: 24.088
alignment_block:
US-09-528-682-1 x AU211806/rev ..
Align seg 1/1 to reverse of: AU211806 from: 1 to: 586
102 GLVAlTySerProHisProtyrGluGlnValSerLeuLeuGIyGI 118
|||||: |||||: : : : : : : : : : : : : : : : : : :
393 GGATTCATCCACCCTACCATCAACAACA...GA 359
118 yileProTyrSerGlnIleTyrGlyTrpTyrArgValAsnPhelYValI 135
358 TCTACCGTATTGGAAAGTGTCTCT.....CACTGAGGAAGA 321
135 leAspeluarGleuHisArgAsnArgGluTyrArgspargTYrTYrAg 151
320 AAGAACAGAGATCCACCGAGTTGCATCTTACCCAAGACATCACCCC 271
152 AsnLeuAsnIlelealProAlagIuaepGIyTYrArgLeuAlaelyphe 168
270 AATGGAGAGATTCCCAAGATACGG.....TATCGTAACCAAGACTA.. 229
168 opRoasPHISglnAlaTRP.....ArgGluGluProTPRIghISH 182
228 ....CATTCAATGCTTGCGAGCCGCTTCGAGACAAGAAGCCTTATC 183
182 tsAlaproGInGIyCysGIyAsnSerSeraTgrThrlleThrgIyaSPthr 198
182 ACCCTCCGCAAGTCGCTCATCACCA..... 157
199 CyasnGluGluThrGlnAsnLeuSerThrlIeTyrLeuArgGLuyrGI 215
156 ...AACCAAGAGATGCGCCA.....CGAACAAGATA 128

```

```

215 nserlrvsVLLysArgGlnIlePheSerAspYrGlnSerGIuValAspI 232
| : : : : : | | | : : : | | | : : : | | | : : : | | | : : : |
127 ACCTCAAGTGATCGACACCGATTCCAAAGACCGACGACGCTTCACAA 78
232 leryTsnArG 235
: : : : : | | |
77 ACCACAGCCGA 67

seq_name: gb_est1:AU221819

seq_documentation_block:
LOCUS AU221819 599 bp mRNA linear EST 17-JUL-2001
DEFINITION AU221819 unpublished oligo-capped cDNA library, stage I1
ACCESSION AU221819
VERSION AU221819.1 GI:14859976
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea
; Rhabditiidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 599)
AUTHORS Kohara,Y., ShIn-1,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..599
Location/Qualifiers
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk898e07"
/clone_1lb="unpublished oligo-capped cDNA library, stage
I1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="I1"
BASE COUNT 117 a 142 c 172 g 168 t
ORIGIN
alignment_scores:
Quality: 80.50 Length: 137
Ratio: 1.184 Gaps: 7
Percent Similarity: 49.635 Percent Identity: 24.088

Alignment_block:
US-09-528-682-1 x AU221819/rev ..
Align seg 1/1 to reverse of: AU221819 from: 1 to: 599

102 gIyValTyrSerProHisProTyrGlnGluValSerAlaLeuGlycI 118
| : : : : | | | | : : : : | | | : : : | | | : : : | | | : : : |
395 GGGATTCCATCACCGTACCATCATCAACAACAA.....GA 361

118 yIleProTyrSerGlnIleTyrGlyrPrpYrArGValaAnPheGlyValI 135
: : : : : | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
360 TCTACCGTATTGGAAAGTCTGCTCT.....CACTGAGGAAGA 323

135 lAspGIuArGLeuHisArGAsnArGILuTyrArGAspArGrYrYrArG 151
: : : : : | | | | : : : : : | | | : : : : : : : : : : : : : : : :
322 AGAACAAGAGATCCACGACGATTCATCTTACCAGAAAGACCATCAGCC 273

152 AsnLeuAsnIleAlaProAlaGluAspGlyTyrArGLeuAlaGlyPhePr 168
| | : : : : | | | | : : : : | | | | : : : : | | | : : : : |

```



```

80  tyrSerThrTyrTyrLe.....TyrValIleLeI 89
117 TMTTGGCGCTGATTTGAAACAGCGTCGATCGAAGCGCAGCAGCGCA 166
100 ..ValIleuGIValTyrSerProHisProTyrGIuGIuGIuValSerAla 115
167 AACCTGCTGGCGCTTTTAAACAAGCGCTGTAAGACAGAGGTGGCAGTGC 216
116 LeuGI.....GlyIleProTyr...SerGIuIleTyr 125
217 AACGCGAGTACTATTGGCCGATGGCACAAAGTTTGTTCATGTGTGA 266
125 r.....GlyT 127
267 CAACAGATCTATTGCGACGACGACACAGTCAACAGCGCAGCAGCT 316
127 rPTyrArgValAsnPhenGlyValIleAspGIuArgLeuHisArgAsnArg 143
317 CGCATGATTTAAACAGCCAGGTAGTGCACAAAGCGCGCGCTCTTTTA 366
144 GIuTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGIuAs 160
367 GAT.....TTCCTCGTTTGA 383
160 pGIuTyrArgLeuAlaGlyPheProProAspHisGlnAlaTTPArgGIu 177
384 CGGATGTGTTTGGCATCGTCGCCCATACACACACCGCAGTAATGAAG 433
177 Lu 177
434 AA 435

seq_name: gb_est1.AU215966

seq_documentation_block:
LOCUS AU215966 621 bp mRNA linear EST 17-JUL-2001
DEFINITION AU215966 unpublished oligo-capped cDNA library, stage I1
ACCESSION AU215966
VERSION AU215966
KEYWORDS AU215966.1 GI:14854123
SOURCE EST.
ORGANISM Caenorhabditis elegans.
REFERENCE Caenorhabditis elegans.
AUTHORS Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae
1 (bases 1 to 621)
1 Rhabditidae; Peloderinae; Caenorhabditis.
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1..621
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/cclone="yK832b05"
/cclone_lib="unpublished oligo-capped cDNA library, stage
I1"
/sex="hermaphrodite"
/rissue_type="whole animal"
/dev_stage="I1"

```


[illegible]

ORGANISM Caenorhabditis elegans

EXPERT ON

2000

COMMENT

COMMENT:

FEATURES

Source

1. .639

/organism="Caenorhabditis elegans"

```
/strain="N2"
```

/db_xref="taxon:6239"

```
/clone="yk870h06"
```

```
/clone_lib="unpublished oligo-capped cDNA library, stage
```

11

```
/sex="Hermaphrodite"
```

/tissue type="whole animal"

```
/dev stage="L1"
```

BASE COUNT	127 a	150 c	179 d	183 f
------------	-------	-------	-------	-------

ORIGIN

alignment_scores:

Quality: 80.50

Quality: 80.50

Ratio: 1.184

similarity: 49.6335

similarity: 49.6335

similarity: 49.6335

alignment_block:

US-09-528-682-1 x AU218912/rev

Align seg 1/1 to reverse of: AU218912 from: 1 to: 639

102 G l y A l a T y s e r P r o h i s t P r o I n g l u t i n S e r L a l e u g l y
 118 | | | | | : : : : :
 409 G G A T T C C A T C A C C G R A C C A T C A T C A A C A C A GA 3755
 118 y l l e P r o T y S e r g i n l l e T y g l Y t r p T y a r g V a l a s n P h e g l y a l l 1355
 374 T C A C C G T T T T T G G A A A G T C T C C T C A C T G A G A A G A 3377
 135 l e a s p g l u a r t g l e u n i s t a d a s n a r g l u t y r a r g a s p a r g t y t y a r g 1515
 336 A A G A C A C A G G A T C C A C C A G A T T G A T C T T A C C C A G A A C C A T C A C C C 2877
 152 A s n l e u a s n l l e a l a p r o a l a g l u s p c l y t y r a r g l e u a l a g l y p h e r 1688
 286 A A T G G A G A G A T T C C C A A G A T A C G T A T C G T A C A C A G A C T A . 2455
 168 o p r o a s p h i s g l a t r p A r g g l u g l u t r o t r i l l e i s h 1822
 244 C A T C A G C T C C T G G A G C G T T C G G A C C A A A A C C G T T A T C 1999
 182 I s a l a p r o g i n g l y c s g l y a s e r s e r a r g t n l l e t h g l y a s p t h r 1988
 198 A C C T C C G C A A G T C G T C A T C A C C A 1737
 199 C y s A s n g l u g l u t h g l i n a s n l e u s e r t h r l l e t y r l e u a r g g l u t y r g l 2155
 172 A A C C A A G A G A G T C G C C A C G A A A G A T C A 1444
 215 n s e r t y s v a l y s a r g g l n l l e p h e s e r a s p t y r g l n s e r g l u a l a s p i 2222
 143 A C C T C A A G T G A T C G A C A C C A G T T C A A A G A C C G G A C C A G G T G C T T C A A 94
 332 l e f y r a s n a r g 235

/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-739-167-1	70.50	91.42	2.1e+03	16566	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-568-472-1	68.50	70.09	3.2e+04	6
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-404-796-1	70.50	91.42	2.1e+03	16566	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-443-501A-2	68.50	69.54	3.8e+04	7
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-931-869-1	70.50	91.42	2.1e+03	16566	/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-09-036-987A-1	68.50	68.26	4.0e+04	8
/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-350-399-1	70.50	91.42	2.1e+03	16566	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-370-700-1	68.50	68.26	4.0e+04	80
/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-730-379-17	70.50	82.53	6.6e+03	35100	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-936-135-3	68.00	117.78	71.49	13
/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-757-669A-17	70.50	82.53	6.6e+03	35100	/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-421-750A-10	68.00	115.76	92.68	16
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-230-371A-17	70.50	79.07	1.0e+04	46899	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-591-468-6	68.00	113.39	125.62	1
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-119A-1	70.00	120.20	52.42	1362	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-591-468-6	68.00	113.39	125.62	1
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-374-155A-7	70.00	130.20	52.42	1362	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-591-468-6	68.00	112.39	142.75	1
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-875-396-7	70.00	117.18	77.28	1755	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-591-468-6	68.00	112.39	142.75	1
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-317-785-1	70.00	113.91	117.53	2308	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-591-468-6	68.00	112.39	142.75	1
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-261-086-7	70.00	113.91	117.53	2308	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-746-789A-1	68.00	111.69	156.24	1
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-09-018-760-3	70.00	113.93	120.27	2343	/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-936-135-3	68.00	107.33	273.20	27
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-073-384C-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-936-135-3	68.00	107.33	273.20	27
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-483-043-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-379-523-4	68.00	105.02	367.29	33
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-481-238-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-921-750A-1	68.00	93.82	1.5e+03	85
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-471-066B-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-921-750A-1	68.00	93.82	1.5e+03	85
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-484-956-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-477-231A-1	68.00	92.69	1.8e+03	94
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-57-653-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-473-750A-4	68.00	92.69	1.8e+03	94
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-599-491-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-473-750A-4	68.00	92.69	1.8e+03	94
/cgn2_6/p/odata/2/1na/5b.COMB	seq:US-08-756-386-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-09-306-987A-1	68.00	67.15	4.6e+04	80
/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-823-516-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-370-700-1	68.00	67.15	4.6e+04	80
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-682-853A-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-09-370-700-1	68.00	67.15	4.6e+04	80
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/6b.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	127.42	2
/cgn2_6/p/odata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/odata/2/1na/5a.COMB	seq:US-08-471-251A-1	67.50	113.28	12	

[illegible]

cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-1801-283A-4	65.00	83.48	5.8e+03	1171.11
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-1922-248-4	65.00	83.48	5.8e+03	1171.11
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-592-874-1	65.00	72.74	2.3e+04	2880.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-09-096-942-2	65.00	72.74	2.3e+04	2880.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-09-096-942-2	65.00	72.74	2.3e+04	2880.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-9321-873A-11	64.50	121.23	45.38	4444.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-998-416-609	64.50	113.45	124.68	867
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-632-434-7	64.50	112.46	141.56	942
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-633-434-5	64.50	112.19	146.42	963
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-632-434-5	64.50	105.57	342.25	1677
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-463-202-25	64.50	103.47	448.16	2004.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-468-434C-32	64.50	103.47	448.16	2004.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-476-316-7	64.50	101.54	574.39	2252
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-461-823-1	64.50	99.78	719.95	2726
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-183-253-1	64.50	99.09	786.05	2887
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-176-258-8	64.50	97.91	914.02	3186
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-477-306-8	64.50	97.52	961.43	3299
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-477-316-41	64.50	97.34	983.41	3342
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-254-989-1	64.50	95.99	1.2e+03	3744.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-348-333-16	64.50	95.99	1.2e+03	3744.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-465-965-16	64.50	95.99	1.2e+03	3744.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-465-966-16	64.50	95.99	1.2e+03	3744.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-784-649A-5	64.50	94.44	1.4e+03	4264.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-784-649A-5	64.50	93.41	1.6e+03	4466.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-181-471-2	64.50	93.35	1.6e+03	4669.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-752-447-1	64.50	93.35	1.6e+03	4669.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-955-957A-1	64.50	92.90	1.7e+03	4844.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-955-957A-1	64.50	92.90	1.7e+03	4844.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-955-957A-6	64.50	89.40	2.7e+03	6505.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-793-610-5	64.50	85.80	4.3e+03	8789.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-793-610-5	64.50	85.80	4.3e+03	8789.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-335-700-2	64.50	84.10	5.4e+03	10133.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-335-700-2	64.50	84.10	5.4e+03	10133.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-924-847A-3	64.00	108.23	243.52	1222.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-924-847A-3	64.00	108.23	243.52	1222.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-09-120-052-1	64.00	108.23	243.52	1224.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-09-120-052-1	64.00	108.23	243.52	1224.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-09-357-251-17	64.00	106.37	308.98	1433.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-09-357-251-17	64.00	101.10	607.45	2224.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-731-965-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-731-965-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-728-975-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-728-975-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-07-917-320-1	64.00	95.99	1.2e+03	3414.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-07-917-320-1	64.00	95.99	1.2e+03	3414.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-888-497-21	64.00	93.16	1.7e+03	4322.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-888-497-21	64.00	93.16	1.7e+03	4322.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-888-497-21	64.00	93.16	1.7e+03	4322.00
cgcn2.6/prodata/2/lna/5B.COMB.seq:US-08-888-497-21	64.00	93.16	1.7e+03	4322.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00
cgcn2.6/prodata/2/lna/5A.COMB.seq:US-08-809-599A-4	64.00	99.34	761.09	2577.00


```

/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-232-200-101 + 63.50 100.31 672.11 2166
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-232-197-101 + 63.50 100.31 672.11 2166
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-232-201-101 + 63.50 100.31 672.11 2166
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-286-767-1 + 63.50 99.50 746.13 2319
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-090-526-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-486-715-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-486-719-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-271-354-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-476-100-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-565-861-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-475-749-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:PCPT-US94-07658-1 + 63.50 99.49 746.62 2330
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-577-184-5 + 63.50 99.41 755.52 2336
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-332-312-1 + 63.50 98.88 807.05 2441
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-446-794-3 + 63.50 97.23 997.84 2804
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-693-697-32 + 63.50 96.92 1.0e+03 2877
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-588-526-1 + 63.50 96.91 1.0e+03 2880
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-355-888-6 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-588-190-1 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-693-697-6 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-640-389-1 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-693-696-6 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-618-957-1 + 63.50 96.46 1.1e+03 2991
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-837-635-6 + 63.50 94.15 1.5e+03 3629
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-09-599-455-3 + 63.50 93.38 1.6e+03 3871
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-09-069-781-3 + 63.50 93.38 1.6e+03 3871
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-459-448-18 - 63.50 90.77 2.3e+03 4817
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-459-555-18 - 63.50 90.77 2.3e+03 4817
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-459-504-18 - 63.50 90.77 2.3e+03 4817
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-459-444-18 - 63.50 90.77 2.3e+03 4817
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-09-547-422-18 - 63.50 90.77 2.3e+03 4817
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-854-585-1 + 63.50 90.05 2.5e+03 5117
/cgn2_6/ptodata/2/1na/6A.COMB.seq:PCPT-US95-0512-1 + 63.50 90.05 2.5e+03 5117
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-439-313-535 + 63.50 87.99 3.3e+03 6082
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-439-313-536 + 63.50 87.99 3.3e+03 6140
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-387-942-16 + 63.50 79.31 9.9e+03 12588
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-390-878-16 + 63.50 75.80 1.5e+04 18885
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-318-448-11 - 63.50 74.65 1.8e+04 18596
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-923-137-2 + 63.50 66.60 4.9e+04 36519
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-161-271-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-449-315-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-444-803-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-457-262-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-456-262-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-456-240-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-455-736-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-971-217-44 + 63.50 112.30 144.40 723
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-065-104-2 - 63.50 106.92 287.97 1135
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-065-104-2 - 63.50 106.92 287.97 1135
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-090-523-9 + 63.50 103.44 449.85 1519
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-398-627-9 + 63.50 103.44 449.85 1519
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-406-858-9 + 63.50 103.44 449.85 1519
/cgn2_6/ptodata/2/1na/5A.COMB.seq:PCPT-US91-04036-9 + 63.50 103.44 449.85 1519
/cgn2_6/ptodata/2/1na/5A.COMB.seq:PCPT-US94-05775-9 + 63.50 102.61 500.19 1628
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-286-691-3 + 63.50 102.61 500.19 1628
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-687-147-3 + 63.50 101.77 557.23 1747
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-202-389-1 + 63.50 101.77 557.23 1747
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-466-589-3 + 63.50 101.71 561.63 1756
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-700-636-3 + 63.50 101.71 561.63 1756
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-467-574-3 + 63.50 101.71 561.63 1756
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-09-217-345-3 + 63.50 101.71 561.63 1756
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-178-252-26 + 63.50 100.37 667.12 1965
/cgn2_6/ptodata/2/1na/6B.COMB.seq:PCPT-US91-05160-1 + 63.50 98.85 810.77 2232
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-070-356-1 + 63.50 98.39 859.64 2319
/cgn2_6/ptodata/2/1na/5A.COMB.seq:PCPT-US92-08090-1 - 63.50 97.66 943.85 2465
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-454-097-34 + 63.50 93.46 1.6e+03 3505

```


[illegible]

cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-180-195-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-477-329-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-475-458-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-980-400-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-583-459A-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-583-210-35	61.50	88.41	3.1e+03	4054	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-168-917-3	61.50	88.28	3.1e+03	4100	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-460-510-3	61.50	88.28	3.1e+03	4100	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-460-490-3	61.50	88.28	3.1e+03	4100	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:PCPT-US92-00070-3	61.50	88.28	3.1e+03	4100	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-851-089-12	61.50	88.15	3.2e+03	4144	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-851-089-13	61.50	88.15	3.2e+03	4144	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-731-193A-10	61.50	86.38	4.0e+03	4808	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-486-273A-10	61.50	86.38	4.0e+03	4808	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-940-086A-10	61.50	86.38	4.0e+03	4808	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-940-035A-10	61.50	86.38	4.0e+03	4808	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-371-008-1	61.50	84.46	5.1e+03	5648	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-462-728-1	61.50	83.02	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-168-917-5	61.50	83.01	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-460-510-5	61.50	83.01	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-460-490-5	61.50	83.01	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:PCPT-US92-00730-5	61.50	83.01	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:PCPT-US92-00862-5	61.50	83.01	6.2e+03	6375	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-015-973-2	61.50	82.03	7.0e+03	6924	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-448-164-2	61.50	82.03	7.0e+03	6924	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-718-388-8	61.50	71.75	2.6e+04	16382	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-470-179-109	61.00	114.28	112.00	423	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-339-706-1	61.00	108.17	245.33	706	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-366-783-3	61.00	107.21	277.40	765	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-693-147-1	61.00	106.27	313.12	828	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-053-197A-4	61.00	106.22	314.85	831	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-085-761A-4	61.00	106.22	314.85	831	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-553-489-9	61.00	105.81	331.83	860	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-469-667-5	61.00	102.48	508.77	1137	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-469-667-5	61.00	101.75	558.91	1209	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:PCPT-US95-07289-5	61.00	101.75	558.91	1209	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-342-681C-16	61.00	100.46	659.46	1347	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-222-124-10	61.00	99.93	705.72	1408	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-222-124-9	61.00	99.93	705.72	1408	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-842-657A-9	61.00	99.88	710.33	1414	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-09-292-768-5	61.00	99.06	789.45	1515	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-230-388-2	61.00	98.59	837.80	1575	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-531-601-2	61.00	98.44	854.96	1596	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-859-032-2	61.00	98.44	854.96	1596	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-202-056-8	61.00	97.35	982.70	1748	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-776-246-1	61.00	95.46	1.3e+03	2048	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-776-246-1	61.00	95.46	1.3e+03	2048	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-613-009A-4	61.00	95.13	1.3e+03	2106	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-613-009A-4	61.00	95.13	1.3e+03	2106	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-517673-5	61.00	94.70	1.4e+03	2182	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-039-555B-12	61.00	94.65	1.4e+03	2191	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-613-003A-3	61.00	94.32	1.4e+03	2247	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-552-369-1	61.00	94.35	1.5e+03	2254	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-552-369-1	61.00	94.35	1.5e+03	2254	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-537-361E-3	61.00	94.28	1.5e+03	2261	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-537-361E-3	61.00	94.28	1.5e+03	2261	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-931-608A-2	61.00	92.98	1.7e+03	2500	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-08-931-608A-2	61.00	92.98	1.7e+03	2500	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-07-952-817-8	61.00	92.43	1.8e+03	2639	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-07-952-817-8	61.00	92.43	1.8e+03	2639	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-5210025-1	61.00	92.43	1.8e+03	2639	
cgcn2_6/p/odotata/2/ina/6A.COMB	seq:US-09-624-763A-16	61.00	91.78	2.0e+03	2787	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-085-761A-5	61.00	90.44	2.4e+03	3120	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-032-382B-2	61.00	90.25	2.4e+03	3170	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-032-382B-2	61.00	90.25	2.4e+03	3170	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-09-226-012-1	61.00	89.13	2.8e+03	3480	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-931-608A-1	61.00	89.11	2.8e+03	3487	
cgcn2_6/p/odotata/2/ina/6B.COMB	seq:US-08-931-608A-1	61.00	89.11	2.8e+03	3487	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-126-012-3	61.00	87.62	3.4e+03	3950	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-08-159-784-1	61.00	87.38	3.5e+03	4031	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:US-09-342-681C-18	61.00	86.79	3.8e+03	4235	
cgcn2_6/p/odotata/2/ina/PCPTUS.COMB	seq:PCPT-US93-04383-15	61.00	86.43	4.0e+03	4365	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-284-941-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-08-447-642-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5B.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09-236-503-1	61.00	86.32	4.0e+03	4403	
cgcn2_6/p/odotata/2/ina/5A.COMB	seq:US-09					

/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-508-004-2	60.50	80.37	8.6e+03	6611	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-481-130-27	60.00	88.99	2.9e+03	2930
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-402-066-2	60.50	80.37	8.6e+03	6611	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-485-604-27	60.00	88.99	2.9e+03	2929
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-987-439-2	60.50	80.21	8.8e+03	6702	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-487-595-27	60.00	88.99	2.9e+03	2928
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-570-311-9	60.50	79.87	9.2e+03	6895	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-261-206a-71	60.00	87.54	3.5e+03	3333
/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-353-485-9	60.50	78.36	1.2e+04	7824	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-09-347-488-1	60.00	87.44	3.5e+03	3332
/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-076-011-1	60.50	69.07	3.6e+04	17041	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-566-398-38	60.00	86.73	3.8e+03	3533
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-07-642-734C-3	60.50	67.02	4.7e+04	20235	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-631-814B-3	60.00	85.73	4.4e+03	3844
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-439-009A-3	60.50	67.02	4.7e+04	20235	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-974-549A-4	60.00	85.70	4.4e+03	3855
/cgn2_6/p/ptodata/2/1lna/6A.COMB.seq:US-09-128-155-16	60.50	42.93	5.7e+05	153331	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-07-862-021B-9	60.00	85.18	4.7e+03	4022
/cgn2_6/p/ptodata/2/1lna/6A.COMB.seq:US-09-128-155-17	60.50	41.16	6.3e+05	176373	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-313-288B-9	60.00	85.18	4.7e+03	4021
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-470-179-195	60.00	111.96	150.83	427	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-566-398-23	60.00	84.24	5.3e+03	4355
/cgn2_6/p/ptodata/2/1lna/6A.COMB.seq:US-08-188-930-223	60.00	108.94	222.21	550	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-332-459-21	60.00	84.22	5.3e+03	4366
/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-946-021A-4	60.00	108.62	230.93	564	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-392-459-25	60.00	84.22	5.3e+03	4367
/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-938-416-730	60.00	107.23	276.88	635	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-392-459-25	60.00	84.22	5.3e+03	4368
/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-938-416-730	60.00	105.88	329.19	711	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-392-459-25	60.00	84.22	5.3e+03	4369
/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-245-808-7	60.00	104.54	390.55	795	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-392-459-25	60.00	84.22	5.3e+03	4370
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-287-442-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/6B.COMB.seq:US-08-392-459-25	60.00	84.22	5.3e+03	4371
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-459-701-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-566-398-35	60.00	80.24	8.8e+03	6099
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-459-701-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/5B.COMB.seq:US-08-566-398-35	60.00	80.24	8.8e+03	6098
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-459-701-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-612-321-3	60.00	80.14	8.9e+03	6143
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-459-701-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-612-321-3	60.00	80.14	8.9e+03	6144
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-459-701-4	60.00	102.01	540.48	983	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-427-640-1	60.00	101.05	611.00	1065	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	7360
/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-137-116-2	60.00	101.02	613.63	1068	/cgn2_6/p/ptodata/2/1lna/5A.COMB.seq:US-08-266-398-32	60.00	77.98	1.2e+04	


```

/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-446-794A-1 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-750-007-2 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-945-024-1 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/1na/Backfiles1.seq:5244792-1 + 59.50 85.85 4.3e+03 3472 1 Pat
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-630-915A-33 + 59.50 83.89 5.5e+03 4091 1
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-09-586-719-11 - 59.50 83.58 5.7e+03 4198 1
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-004-139B-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-811-492-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/1na/PCTUS.COMB.seq:PCT-US96-10545A-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-436-332B-9 + 59.50 81.84 7.2e+03 4858 1
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-09-073-354-6 + 59.50 80.71 8.3e+03 5339 1

```

```
seq_name: /cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-823-120-5
```

```

seq documentation block:
; Sequence 5, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Piza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,003
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..711
; US-08-823-120-5

```

```

alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333

```

```
alignment_block:
US-09-528-682-1 x US-08-823-120-5 ..
```

```
Align seg 1/1 to: US-08-823-120-5 from: 1 to: 711
```

```

seq_name: /cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-823-120-7
seq documentation block:
; Sequence 7, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Piza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines

```



```

1 AsnGlyAspLeuLeuTYRArgAlaAspSerArgProProAsnGluIleY 17
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 AATGATGATATAAGTATTATCGGGCAGATTCTGACCTCCGATGAAATMAA 50
17 sArgSerGlyGlyLeuLeuMetProArgGlyLHisAsnGluTYRpheAspArg 34
51 GCGATCGAGGGGTGCTTATGCGCAAGGAGGACAGAGTAGTCTTTGACCAG 100
34 LyrHisGlnMetAsnIleAsnLeuTYRAspHisAlaArgGlyThrGlnThr 50
101 GTCATCAAAATGAATATCAACCTTTATGATCATGCAAGAGGAATCAGAGC 150
51 GlyPheValArgTYRAspAspGlyTYRValSerThrSerLeuSerLeuArg 67
151 GGATTGTGTAGGCCAGATGATGGAATATGTTCCACCTCAATATGATTGGAG 200
67 gSerAlaHisIleuLeuIleGlyGlnSerIleLeuSerGlyTYRSerThrTYR 84
201 AAGGCCCACTTATGCGTCAACAATATGTCTGCTCATTTCTACTATTAT 250
84 yIleIleTYRValIleAlaIleThrAlaProAsnMetPheAsnValAsnAspVal 100
101 ATATTATATTGTTATAGCCACTCACCACCAACATGTTTAAACCTTATATGTTGA 300

```

```

1      seq.documentation_block:
2      Sequence 1, Application US/08435605A
3      Patent No. 5874287
4      GENERAL INFORMATION:
5      APPLICANT: Burnette, W. Neal
6      APPLICANT: Kailow, Harvey R.
7      TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
8      TITLE OF INVENTION: SUBUNIT ANALOGS
9      NUMBER OF SEQUENCES: 57
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Amgen Inc.
12     STREET: 1840 De Havilland Drive
13     CITY: Thousand Oaks
14     STATE: California
15     COUNTRY: USA
16     ZIP: 91320-1789
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: PatentIn Release #1.0, Version #1.30
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/435,605A
24     FILING DATE: 05-MAY-1995
25     CLASSIFICATION: 435
26     ATTORNEY/AGENT INFORMATION:
27     NAME: Mazza, Richard J.
28     REGISTRATION NUMBER: 27,657
29     REFERENCE/DOCKET NUMBER: A-196B
30     INFORMATION FOR SEQ. ID NO: 1:
31     SEQUENCE CHARACTERISTICS:
32     LENGTH: 774 base pairs
33     TYPE: nucleic acid
34     STRANDEDNESS: single

```


TOPLOGY: linear
MOLECULE TYPE: CDNA
US-08-435-605A-1

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-435-605A-1 ..

Align seg 1/1 to: US-08-435-605A-1 from: 1 to: 774

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGATAGATTATATCGGCAGATTCTAGACCTCTGATGAATAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||
105 GCAGTCAGGCTGCTTATGCCAAGACAGAGTAGTACTTGACCGAG 154
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGACG 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
205 GGATTTGTTGGCAGCATGATGATATGTTCCACCTCATTTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
255 AAGTGGCCACTAGTGGTCAACTATATGTCGTGATCATCTATAT 304
84 YrLleTyrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
305 ATATATATGTTATAGCCACTGACCCCAACATGTTTAACTTAATGATGA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
355 TTAGGGGCATACAGTCTCTCATCCAGATGAAACAAGAACTTCTGCTTAG 404
117 YgIlyIleProTyrSerGlnIleTyrGlyTyrTyrValAsnPheGly 134
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
405 TGGGATTTCCATCTCCCAATATATGATGATGATCGATTCATTTGGGG 454
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
455 TGGTGTGATGACAATTACATCGTAATAGGGGCTACAGAGATGATATAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
505 AGTAACTTATGATATGCTCCAGCAGCAGATGTTATGATGATGACAGTTT 554
167 eRProAspHisGlnAlaTatPArgGluGluProTyrIleHisAlaP 184
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
555 CCTCCGAGACATAGAGCTTGAGGAGAGAGCGGTGATTCATCATGAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
605 CGCGGGGTTGTGGAAATGCTCAAGATCATGATGATTAATCTTGGCAT 654
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
655 GAAAAAACCAAGCTAGTGTAAATTCCTTGACGAATACCAATCTAA 704
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIle 234
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
705 AGTTAAAGCAATATATTTTACAGGATATCATCTGATATTGATACACA 754
234 snArgIleArgAspGluLeu 240
|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```

755 ATAGAAATTAAGATGAATTA 774

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-449-045C-1

seq_documentation_block:

```

; Sequence 1, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449, 045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..777
; US-08-449-045C-1

```

alignment_scores:

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-449-045C-1 ..

Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGATAGATTATATCGGCAGATTCTAGACCTCTGATGAATAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||
105 GCAGTCAGGCTGCTTATGCCAAGACAGAGTAGTACTTGACCGAG 154
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGACG 204

```


Align seg 1/1 to: 5223610-1 from: 1 to:20220

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProDroAspGluIleLy 17
111 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
570 AATGATGATAAGTATATATCGCGCAGATTTCTAGACCTCTGATGAAATAA 619
17 sarGSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
1:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
620 GGAGTCAGGTGGTCCATGTCACCAAGAGACAGAGAGTACTTTGACCAG 669
34 IYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
1111111111111111111111111111111111111111111111111
670 GTRACTAAATGAATATCAACCTTTATGATATATGCAAGAGACACTCAGACG 719
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuArg 67
720 GGATTTGTTAGGACACGATGATGATATGTTTCCACCTCAATTAAGTTGAG 769
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
1111111111111111111111111111111111111111111111111
770 AAGTGGCCACTTAAAGTGGGTCAACATAATTTGTGCGTCATCTCACTATT 819
84 TyrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAspAspVal 100
820 ATATATATGTTATATAGCCACTGACCCACACATGTTTAAAGCTTAAATGATGA 869
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
1111111111111111111111111111111111111111111111111
870 TTAGGGGCGATACAGTCTCTCATCCAGATGAACAACAGTTTCTGCTTATAGG 919
117 yGlyIleProTyrSerGlnIleTyrGlyTTPtyrArgValAsnPheGly 134
1111111111111111111111111111111111111111111111111
920 TGGGATTCATACTCTCCCAATATATGATGTGATCGAGTTCAATTTTGGGG 969
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrCyr 150
970 TGGTATGATGACAAATATACATCGTAATATGGGGCTACAGAGATGATATATAC 1019
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
1111111111111111111111111111111111111111111111111
1020 AGTAACCTTAGATATGTGCTCCAGCAGCAGATGGTATGATTTGGCAGGTTT 1066
167 eProProAspHisGlnAlaTTPArgGluGluProTyrIleHisHisAlaP 184
1111111111111111111111111111111111111111111111111
1070 CCTCTCCGGAGCAATAGAGCTTGGAGGGAGAGCCCTGTGATTCATCATGAC 1119
184 roGlnIleCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
1111111111111111111111111111111111111111111111111
1120 CGCGGGTGTGGGAATGCTCCAGACATCATGATGAGTAATATCTTGGCAT 1166
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
1111111111111111111111111111111111111111111111111
1170 GAAATAAACCAAGTCTAGGTGTAATAATTCCTTCACGATATCCCAATCTAA 1219
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
1111111111111111111111111111111111111111111111111
1220 AGTTAAAGCAATATATTTTCAGCGCTATCAATCGATATGATACACATTA 1266
234 snArgIleLeuAspGluLeu 240
1111111111111111111111111111111111111111111111111
1270 ATRGAAATTAAGGATGAAATTA 1289
```

seq_name: /cgn2-6/ptodata/2/ina/5B.COMB.seq:US-08-435-605A-5

```

seq_documentation_block
: Sequence 5, Application US/08435605A
: Patent No. 5874287
:
: GENERAL INFORMATION:
:
: APPLICANT: Burnette, W. Neal
:
: APPLICANT: Kaslow, Harvey R.
:
: TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXININ
:
: TITLE OF INVENTION: SUBUNIT ANALOGS
:
: NUMBER OF SEQUENCES: 57

```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 582 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-435-605A-5

```

```

alignment_scores:
Quality: 934.00 Length: 193
Ratio: 5.049 Gaps: 0
Percent Similarity: 95.855 Percent Identity: 86.528

```

```

alignment_block:
US-09-528-682-1 x US-08-435-605A-5 ..
Align seg 1/1 to: US-08-435-605A-5 from: 1 to: 582

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
1 AATGATGATTAAGTATATATCGGCAATCTCTAGACCTCGATGAATAAA 50
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
1:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 51
GCAGTCAGGTGCTTATGCCAAGAGACAGAGTAGTACTTTGACCGAG 100
34 lYThGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 101
GTACTCAATGAATATCAACCTTTATGATCATGCAAGAGAACTCAGACG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 151
GGATTGTGTTAGGACGATGATGATGTTTCCACCTCAATTAGTTTGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 201
AAGTGCCTCCTAGTAGGTGCAACATATATGCTGCTACTTACTACTAT 250
84 yrlEtyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 251
ATATATATGTTATAGCCACCTGACCCCAACATGTTAACTTATATGTA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301
TTAGGGGCAATACAGTCTCATCAATGAACAAGAGTTTGTGCTTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 351
TGGATTCATCTCTCCCAATATATGATGATGATGATGATGATGATGAT 400
134 allEAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 401
TCTTGATGACAAATTAATCTATAGGCGCTACAGAGATGATATTAAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 451
AGTAACTTAGATATGCTCTCCAGCAGACAGATGTTATGATTTGGCAGTT 500
167 eProProAspHisGlnAlaTyrPArgGluGluProThrPheHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 501
CCCTCCGGAGCATAGAGCTTGAGGAGAGAGCCGTGATTCATCATGAC 550
184 roGlnGlyCysGlyAsnSerSerArgThr 193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 551
CGCCGGGTGTGGGAATGCTCCAGATCA 579

```

```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-435-605A-7

```

```

seq_documentation_block:
Sequence 7, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
TITLE OF INVENTION: SUBUNIT ANALOGS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-435-605A-7

```

```

alignment_scores:
Quality: 933.00 Length: 192
Ratio: 5.071 Gaps: 0
Percent Similarity: 95.833 Percent Identity: 86.979

```

```

alignment_block:
US-09-528-682-1 x US-08-435-605A-7 ..

```

```

Align seg 1/1 to: US-08-435-605A-7 from: 1 to: 576

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
1 AATGATGATTAAGTATATATCGGCAATCTCTAGACCTCGATGAATAAA 50
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
1:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 51
GCAGTCAGGTGCTTATGCCAAGAGACAGAGTAGTACTTTGACCGAG 100

```



```

34  LyrhrglnmetasnlleasleuTyTAspHisAlaArgGlyThrGlnThr 50
|||||
101 GTACTCAAAATATTCACCTTATGATCATCCAGAGACCTGACAGC 150
51  GlyPheValArgTyTAspArgGlyTyTValSerThrSerLeuSerLeuArg 67
|||||
151 GCATTTGTTGGCAGCATGATGATGTTCCACCTCAATTAAGTTTGAG 200
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyTSerThrTyT 84
|||||
201 AAGTGCACCATGTAGTGGTCAACATATATGTCTGTCATCTACTATAT 250
84  yTleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
251 ATATATATGTTATAGCCATGACATGCCAACATGTTTAAAGTTAATGAGTA 300
101 LeuGlyValTyTSerProHisProTyTArgGlnGluValSerAlaLeuGln 117
|||||
301 TTAGGGGCATACAGTCTCTCATCCAGATGACAAAGAACTTCTGCTTAGG 350
117 yGlyIleProTyTSerGlnIleTyTArgGlyTyTArgValAsnProGly 134
|||||
351 TGGGATTCCTACTCCCAATATATGATGATGATGATGATGATGATGATG 400
134 alIleAspArgLeuIleAsnArgGlyTyTArgAspArgTyTArg 150
|||||
401 TGCTTGATGACATATACATGATGATGATGATGATGATGATGATGATGAT 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyTArgLeuAlaGly 167
|||||
451 AGTAACTTAATATGCTCCAGCAGCAGATGTTATGATGATGATGATGATG 500
167 eProProAspHisGlnAlaThrArgGlnGluProTyTArgIleHisAla 184
|||||
501 CCTCCGAGCATAGAGCTTGAGAGGAAGAGCGTGATTCATCATGATGCAC 550
184 roGlnGlyCysGlyAsnSerSerArg 192
|||||
551 CGCCGGGTGTGGAAATGCTCAAGA 576
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:us-08-435-605A-9

seq_documentation_block:
; Sequence 9, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-1968
; INFORMATION FOR SEQ ID NO: 9:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-435-605A-9

alignment_scores:
Quality: 157.00 Length: 44
Ratio: 3.925 Gaps: 0
Percent Similarity: 90.909 Percent Identity: 65.909

alignment_block:
US-09-528-682-1 x US-08-435-605A-9 ..

Align seg 1/1 to: US-08-435-605A-9 from: 1 to: 138

197 AspHrcysAsnGluGluThrGlnAsnLeuSerThrIleTyTLeuArgGln 213
|||||
7 AATACCTGCGATGAAAAACCAAGCTAGGTAAATAATCCCTTGACGA 56
213 uTyTArgSerTyValAspArgGlnIlePheSerAspTyTArgGlnSerGlu 230
|||||
57 ATACCAATCTAAAGCTTAAAGACAAATATTTTCAGGCTATCATCTGATA 106
230 alAspIleTyTAsnArgIleArgAspGluLeu 240
|||||
107 TTGATACACATTAATAGAAATTAAGATGATGAATTA 138
seq_name: /cgn2_6/ptodata/2/lna/backfiles1.seq:5223610-4

seq_documentation_block:
; Patent No. 5223610
; APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
; TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
; HORMONE PROMOTER
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,852
; FILING DATE: 18-MAY-1990
; SEQ ID NO: 4
; LENGTH: 4207
5223610-4

alignment_scores:
Quality: 90.50 Length: 338
Ratio: 0.730 Gaps: 16
Percent Similarity: 36.686 Percent Identity: 20.710

alignment_block:
US-09-528-682-1 x 5223610-4 ..

Align seg 1/1 to: 5223610-4 from: 1 to: 4147

5 LeuTyTArgAlaAspSerArgProProAspGluIleLeuArgSerGlyGln 21
|||||
627 GTATACCGCTATGACTCCGCCGCCGAGGAGACTTTCCAGAAC...GG 673
21 yLeuMetProArgGlyHisAsnGluTyTArgPheAspArgGlyThrGlnMeta 38
|||||
674 ATTCACGGCGTGGGAAACAGCAGC..... 698
38 snIleAsnLeuTyTAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
699 ....AATGCTCGACCATCTGACCGGAGCTTCTCCAGTGGCAGC 743
55 TyTAspAspGlyTyTValSerThrSerLeuSerLeuArg..... 67
|||||
744 AGCAACAGCGCTTTCGTCTCCACGACGAGCGCGCTATACGAGGT 793

```



```

68 .....SerAlaHisLeuAlaGlyG 74
794 CTAATCTCGAACAATCGCATCGACGAAGCGGTGCGAGCGCCGAACGCGCGCA 843
74 InsertLeuSerGlyTyrSerThrTyrTyrIleTyrValIle..... 88
      ::::: ::::: ::::: ::::: ::::: :::::
844 GGGGC.....ACGGGCACTTATCGTATCGATCTAAGATCGGGGCC 887
89 .....AlaThrAlaProAsnMetPheAsnValAsnAspVala 100
888 GACAAACAATTCTACGGCGCGCGAGCTCTACTTTCGAATAACGTGCACAC 937
100 IleuGlyValTyrSerProHis.....ProTyrG 110
      ::::: ::::: ::::: ::::: :::::
938 TTAATGGCGACAAATGCGCGCGGTATCTCGCGCGCGCGCTGGCCACTTACC 987
110 IuGInGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly 126
      ::::: ::::: ::::: ::::: :::::
988 AGAGCGCAATATCTCGGACACACCGCGCGCTTCGCCCGCAAAACATCGCAGG 1037
127 TrrTyrArgVal..... 130
      ::::: :::::
1038 GTAACCGGGGTCTATGACAAAGCATACCGCGGAGACCAAGCAACAGGA 1087
131 .....AspPheGlyValIleAspGluAlaGlyLeuHisAlaGln..Arg 144
      ::::: ::::: ::::: ::::: :::::
1088 CTAATTCACACGCTCGCTACGTACGTACGACGACAGACTCGCGCCAAATCCCAAC 1137
144 IuTrrTyrAspArg..... 148
      ::::: :::::
1138 CCAACACATCGCAGAGTCCGTCAGCGTCGATCGTCGACATTTGGTCAT 1187
148 ..... 148
1188 GCGCGCGGTGATAGCGCTTCATGAGCGCGGACAGCCGAAAGCTCCGAGGC 1237
149 .....TyrT 150
1238 CATGCAAGCTGTGTCCGAACGCGCGCGAGCGAGGTCTCGTGTACT 1287
150 Yr..... 150
1288 ACGAAAGCATCGGGTATCTGTTAGACCTGCGCCACCGCCGCAATC 1337
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly.. 166
      ::::: ::::: :::::
1338 CGGTAAATGTGAACAGCATGCCGATGACCGCAAGACGTGCGCCAGGCAT 1387
167 ...PheProProAspHisGlnAlaTrpAlaGlyGluProTrpIleHis 182
      ::::: ::::: :::::
1388 CGTCATTTGCGCGCCAGGACAGCAATTACCCAG.....CATG 1422
182 IsAlaProGlnGly...CysGlyAsnSerSerArgThrIleThr..... 195
      ::::: ::::: ::::: ::::: :::::
1423 GCGAGCCCTATGACATGCGCGCAACAGACCGTGGCCCTGACCGTGGCG 1472
196 .....GlyAspTrpCysAsnGluGluThrGlnAsnLeu 207
      ::::: :::::
1473 GAATTCGCGCGCAGCGCGCAT..... 1493
207 rTrIleTyrLeuArgGluTyrGlnSerIysValLysArg.....GlnI 222
      ::::: ::::: ::::: :::::
1494 .....CTCCAGGAGTACTGCGTCAATGTGACCGCGCGCTGGTGCACAA 1533
222 IephSerAspTyr 226
      ::::: :::::
1534 TATTTCGCGCTTAC 1547
seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5244657-4
seq document number:
; patient NO. 5244657
; APPLICANT: KLIEN, MICHEL A.;BOUX, HEATHER A.;COCKLE,

```

```

STEPHEN.;LOOSMORE, SHEENA M.;ZEALEY, GAVIN R.
TITLE OF INVENTION: GENETIC DETOXIFICATION OF PERTUSSIS
; TOXIN
;
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,423
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 275,376
; FILING DATE: 23-NOV-1988
; SEQ ID NO:4:
; LENGTH: 4208
5244657-4

```

[illegible]


```
189 AsnSerArgThr.....IleThrGlyAspThrCysasn1 201
::: |||||::: |||||::: |||||
1035 AGGTAACCGGGGTCTATGCACACGGCATCACCGGAGACACACGACAC 1084
201 uGluThrGlnAsn 205
||| ::|||
1085 GGAGTATTCACAC 1097

seq_name: /cgn2_6/ptodata/2/1na/backfile1.seq:5433945-4

seq_documentation_block:
; Patent No. 5433945
; APPLICANT: KLEIN, MICHEL H.; BOUX, HEATHER A.; COCKLE,
; STEPHEN A.; LOOSMORE, SHEENA M.; ZEALEY, GAVIN R.
; TITLE OF INVENTION: IMMUNOPROTECTIVE GENETICALLY-DETOXIFIED
; MUTANTS OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/979,798
; FILING DATE: 20-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 589,423
; FILING DATE: 28-SEP-1990
; APPLICATION NUMBER: 275,376
; FILING DATE: 23-NOV-1988
; SEQ ID NO: 4
; LENGTH: 4208
5433945-4

Alignment_scores:
Quality: 89.00 Length: 221
Ratio: 1.000 Gaps: 11
Percent Similarity: 40.271 Percent Identity: 23.529

Alignment_block:
US-09-528-682-1 x 5433945-4 ..

Align seg 1/1 to: 5433945-4 from: 1 to: 4208

5 LeuTyrArgAlaAspSerArgProAspGluIleLysArgSerGly1 21
::: |||||::: |||||::: |||||::: |||||::: |||||
627 GATACCGCTATGACTCCCGCGCGGAGACGTTTCCAGAC...GG 673
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
::: |||||::: |||||::: |||||
674 ATTCAAGCGCGTGGGAAACACGAC..... 698
38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||::: |||||::: |||||
699 ....AATGTGCTCGAACATCTGACCGGACGTTCTCGCAGTCGGCAGC 743
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArg..... 67
::: |||||::: |||||::: |||||
744 ACCAACAGCGCTTGTCTGTCCACACGACGCGCGCTATACGAGGT 793
68 .....SerAlaHisLeuAlaGlyG 74
794 CTATCTGACATCCATCGACAGAACGCGTCCAGCGCAACCGCGCGCA 843
74 InSerIleLeuSerGlyTyrSerThrTyrIleTyrValIleAlaThr 90
::: |||||::: |||||::: |||||
844 GGGGCG....ACCGGCGACTTCATCGCTACATCTACGAAAGTC...CGC 884
91 AlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerProH1 107
||| ||| |||
885 GCCGACAAACAATTTC..... 899
107 sProTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnI1 124
899 ..... 899
124 l eTyrGlyTyrPtyrArgValAsnPheGlyValIleAspLysArgLeuHis 140
```

```
900 ..TACGGCCCCCGCACCTGTACTTC..... 923
||| |||
141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn....11 155
||| ||| ||| |||
924 .....GAATACGTCGACACTTATGGCAACATGGCGCGTATCTCT 964
155 eAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsnHis 172
||| ||| ||| |||
965 CGCGCGCGCG.....CTGGCAGCTTACACAGAGCAATATTC 999
172 InAlaTyrArgGluGluProTyrIleHisAlaProGlnGlyCysGly 188
||| |||||::: |||
1000 TGGCACACCGCGCGCATTCG.....CCGAAACATCCCG 1034
189 AsnSerArgThr.....IleThrGlyAspThrCysasn1 201
::: |||||::: |||||::: |||||
1035 AGGTAACCGGGGTCTATGCACACGGCATCACCGGAGACACACGACAC 1084
201 uGluThrGlnAsn 205
||| ::|||
1085 GGAGTATTCACAC 1097

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-1

seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007_00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Kv
US-09-103-840A-1

Alignment_scores:
Quality: 94.50 Length: 230
Ratio: 1.086 Gaps: 13
Percent Similarity: 37.826 Percent Identity: 24.348

Alignment_block:
US-09-528-682-1 x US-09-103-840A-1/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-1 from: 1 to: 4411529

6 TyrArgAla.....AspSerArgProAspGluIleLysArgse 19
||| ||||| ||||| |||||
3735564TATCGAGCCCATCCACATCTTCCGACATCCCGATCAACATCAAGAA 3735515
19 r.....GlyGlyL 22
3735514CCCTTACCTCGGACCTTGTGTCGCCGCCCATCAACGTCCCGCAATC 3735465
22 eumetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsn 38
||| ||| ||| ||| ||| |||
3735464TGCCTCGCGGTGGCATYACCAATATTTTCATCGCGCCCATCA... 3735421
39 IleAsnLeuTyrAspHis...AlaArgGlyThrGlnThrGlyPheValAr 54
||| ||| ::| |||||:::
3735420...AATCAATCCCATCACCCCTGTGGCGGACAA... 3735388
```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,848A
FILING DATE: 26-AUG-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5113-046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNTE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-296-848A-1

alignment_scores:
Quality: 86.00 Length: 15
Ratio: 5.733 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-296-848A-1 ..

Align seq 1/1 to: US-08-296-848A-1 from: 1 to: 45

186 G1CysG1yAnSerSerArgThrIleThrG1yAspThrCysAsn 200
1 GGTGTGGAATTCATCAAGACATTCACGTCATCTGTAT 45

seq_name: /cgn2.6/ptdata/2/1na/5A.COMB.seq:US-08-204-656B-9

seq_documentation_block:
Sequence 9, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuo
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Library of chromosomal DNA from Bacillus
LIBRARY: macerans, pMAC, generated by treating chromosomal DNA from Bacillus
LIBRARY: IM1243 with a restriction enzyme, and inserting and linking restric
LIBRARY: fragments to pBR322
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2061
OTHER INFORMATION: /note= "Nucleotides 1-2061
OTHER INFORMATION: correspond to nucleotides 82-2142 of the Bacillus macerans
OTHER INFORMATION: cyclomalodextrin glucanotransferase structural gene"

US-08-204-656B-9

alignment_scores:
Quality: 83.50 Length: 241
Ratio: 0.852 Gaps: 13
Percent Similarity: 40.664 Percent Identity: 21.577

alignment_block:
US-09-528-682-1 x US-08-204-656B-9 ..

Align seq 1/1 to: US-08-204-656B-9 from: 1 to: 2061

18 ArgSerG1yLeuMetProArgG1yHisAsnG1yThrPheAspArgG1 34
391 AGGTGCGATCGATTCGCCGCCACACACAC..... 420
34 yThrG1MetAsnLeuAsnLeuTyAspHisAlaArgG1yThrG1nThrG 51
421GTCTCGGCCGACGACAGGCGACACCCG 448
51 lypheValArg.....TyrAspArgG1yTyValSerThrSer 63
449 GCTTCGCCGACAGCGTCCGCTGTATGATACGCT..... 483
64 leuSerLeuArgSerAlaHisLeuAlaG1yC1nSerIleLeuSerG1yTy 80
484TCGCTGCTCGGCCGCTA 500
80 rSerThyTyrTyrIleTyValIleAlaThrAlaProAsnMetPheAsnV 97
501 CAGC..... 504
97 AlaAsnAspValLeuG1yValTySerProHisProTyG1nG1nVal 113
505 ..AATGATACGCGCGCGCTTTTCATCATAC..... 534
114 SerAlaLeuG1yG1yIleProTySerG1nIleTy...G1yTyPyrArg 129
535GGGGGACCGATTTTCCACAGTGTGACAGCGGTATTTACAA 575
129 gValAsnPhcG1yAlaIleAspArgLeuArgLeuHisArgAsnArgG1yTy 146
576 GAACCTTACGACCTGGCGGAC.....ATCAGCATTAACAACAACGCTA 619
146 rGAspArgTyTyTyArg.....AsnLeuAsnIleAlaProAlaG1u 159
620 TGCAGCGCTTATTAAAGCGCTATCGACCTTGGCTCGGCGATGGGTGTG 669


```

160 AspGlyTyrArgLeuAlaGly..... 166
167 .....PheProAspHisGlnAlaTrpArgGlu 177
720 AAGCTTCGTTCCGATTCAGCGCGGATCCGGTATTTACGTCG 769
177 LuProTrpLeuHisAlaProGlnGlyCysGlyAsnSerArgThr 193
770 GGAATGGATCTTGGCGCGATCAACCGACGAGACACATTAA... 816
194 IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
817 .....TTCGCCAACGAAAGCGGATGAACCTGCTGGACTTT.. 852
210 rLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAsp...T 226
853 .....GAATACGCGCAGGAAGTG...CGGAAAGTGTTCGGGACAAAA 892
226 yrGlnSerGluValAspIleTyr 233
893 CGGAACGATGAGGATCTCTAT 915
seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:us-08-470-702-5

```

```

seq documentation block:
: Sequence 5, Application US/08470702
: Patent No. 5631149
: GENERAL INFORMATION:
: APPLICANT: MATSUI, IKUO
: APPLICANT: ISHIKAWA, KAZUHIKO
: APPLICANT: MIYAIRI, SACHIO
: APPLICANT: HONDA, KOICHI
: TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
: TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
: NUMBER OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22042
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/470,702
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/204,656
: FILING DATE: 02-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: WEINER, MARC S.
: REGISTRATION NUMBER: 32,181
: REFERENCE/DOCKET NUMBER: 234-252P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2061 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (synthetic)

```

```

: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: US-08-470-702-5

```

```

alignment_scores:
  Quality: 83.50      Length: 241
  Ratio: 0.852       Gaps: 13
  Percent Similarity: 40.664   Percent Identity: 21.577

```

```

alignment_block:
US-09-528-682-1 x US-08-470-702-5 ..

```

```

Align seg 1/1 to: US-08-470-702-5 from: 1 to: 2061

```

```

18 ArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArgGlu 34
391 AGGTCGGATCGACTTCGCCCAACACAC... 420
34 yThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrG 51
421 .....GTCTCCGGCCGACGACGAGGACGACCCCG 448
51 LyrPheValArg.....TyrAspAspGlyTyrValSerThrSer 63
449 GCTTCGCCGACGACGACGTCGCTGTATGATACGCT..... 483
64 LeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyr 80
484 .....TTCGCTCGCGCGCTA 500
80 rSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnV 97
501 CAGC..... 504
97 AlaAsnAspValLeuGlyValIleTyrSerProHisProTyrGluGlnGluVal 113
505 ..AATGATACGCGCGCGCTTTCATCATAC..... 534
114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyr...GlyTyrTyrArg 129
535 .....GGGGGACCGATTTTCCACGATTAAGACGATATTATTCAA 575
129 gValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyr 146
576 GAACCTGACGACACCTGGCGGAC.....ATCACATTAACACACACCTA 619
146 rGAspArgTyrTyrArg.....AsnLeuAsnIleAlaProAlaGlu 159
620 TGAACGCTTATTATTAAGCGCTATGACACCTTGGTCGCGATGGGTG 669
160 AspGlyTyrArgLeuAlaGly..... 166
670 GACGGGATTCGTTTGAACGGGTGAACAGATATCTTCGCTGGCAAAA 719
167 .....PheProAspHisGlnAlaTrpArgGlu 177
720 AAGCTTCGTTCCGATTCAGCGCGGATCCGGTATTTACGTCG 769
177 LuProTrpLeuHisAlaProGlnGlyCysGlyAsnSerArgThr 193
770 GGAATGGATCTTGGCGCGATCAACCGACGAGACACATTAA... 816
194 IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
817 .....TTCGCCAACGAAAGCGGATGAACCTGCTGGACTTT.. 852
210 rLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAsp...T 226
853 .....GAATACGCGCAGGAAGTG...CGGAAAGTGTTCGGGACAAAA 892
226 yrGlnSerGluValAspIleTyr 233

```


893 CGGAACGATGAAGATCTCTAT 915

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-467-831-5

seq_documentation_block:

; Sequence 5, Application US/08467831
; Patent No. 5635378

; GENERAL INFORMATION:

; APPLICANT: MATSUI, IKUO

; APPLICANT: ISHIKAWA, KAZUHIKO

; APPLICANT: MIYAIRI, SACHIO

; APPLICANT: HONDA, KOICHI

; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

; STREET: 8110 Gatehouse Road, Suite 500 East

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22042

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,831

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/204,656

; FILING DATE: 02-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: WEINER, MARC S.

; REGISTRATION NUMBER: 32,181

; REFERENCE/DOCKET NUMBER: 234-252P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2061 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (synthetic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-467-831-5

alignment_scores:

Quality: 83.50 Length: 241

Ratio: 0.852 Gaps: 13

Percent Similarity: 40.664 Percent Identity: 21.577

alignment_block:

US-09-528-682-1 x US-08-467-831-5 ..

Align seq 1/1 to: US-08-467-831-5 from: 1 to: 2061

```

18 ArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrRheAspArgG1 34
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 AGTCGGATGACTTCCGCCCAACACAC..... 420
34 yThGlmetsnllsasnleuTyrAspHisAlaArgGlyThrGlnThrG 51
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 .....GTCTCCGCCGCGACGAGGACGAAACCCG 448

```

```

51 lYpHeValArg.....TyrAspAspGlyTyrValSerThrSer 63
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
449 GCTTCGCCGAGACGTCGCGTATATACGCT..... 483
64 LeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTy 80
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
484 .....TCGCTGCTCGCGCGCTA 500
80 rSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetHeAsnV 97
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
501 CAGC..... 504
97 alaAsnAspValLeuGlyValTyrSerProHisProGlyGluGlnVal 113
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 ..ATGATACGCGCGCGCTTTCATCATAC..... 534
114 SerAlaLeuGlyIleProTyrSerGlnIleTyr...GlyTyrTyrArg 129
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
535 .....GGGGGACCGATTTTCCACGATGACGCGTATTTACAA 575
129 gValAsnRheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrA 146
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
576 GAACCTCTACGACCTGCGGAC.....ATCACCATTAACAACAACCTA 619
146 rGAspArgTyrTyrArg.....AsnLeuAsnIleAlaProAlaGlu 159
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
620 TGGACGCTTATTTTAAAGCGCTATGACCTTGGCTGGCGATGGGATG 669
160 AspGlyTyrArgLeuAlaGly..... 166
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
670 GACGGGATGCTTTTGAACGGCGTGAAGAGATGCTTTCGCGTGCACAAA 719
167 .....PheProProAspHisGlnAlaThrArgGlu 177
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
720 AAGCTGCTTCCTCGATTACGCGCGGATACCGGATATTATTCAGTTCG 769
177 lUpProPrlLeuHisAlaProGlnGlyCysGlyAsnSerSerArgThr 193
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
770 GGGATGGTATCTTGGCGGATCAACCGACGACACAACTTAA... 816
817 .....TTCGCCAACGAAGCGGATGAACCTGCGACTTT... 852
194 lIeThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTy 210
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 rLeuArgGlyTyrGlnSerLysValLysArgGlnIlePheSerasp...T 226
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
853 .....GAATACCGCGACGAAAGTG...CGGAAAGTGTTCGGGACAAA 892
226 yGlnSerGluValAspIleTyr 233
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
893 CGGAACGATGAAGATCTCTAT 915

```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-195-152-1

seq_documentation_block:

; Sequence 1, Application US/08195152

; Patent No. 5679541

; GENERAL INFORMATION:

; APPLICANT: Bonini, Nancy M.

; APPLICANT: Leiserson, William M.

; TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM: Floppy disk


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,152
; FILING DATE: 14-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Treccartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59551/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-195-152-1

```

```

alignment_scores:
    Quality: 82.50      Length: 269
    Ratio: 0.833       Gaps: 14
    Percent Similarity: 36.803   Percent Identity: 20.818

```

```

alignment_block:
US-09-528-682-1 x US-08-195-152-1 ..

```

```

Align seg 1/1 to: US-08-195-152-1 from: 1 to: 3231

```

```

7  ATGAlAspSerArgProProAspGluIleLysArgSerGlyLysLeu 23
   : : : : : : : : : : : : : : : : : : : : : : : :
1002 AAGTCGGAGTCCCGCAGCCGCGGCAAGTCGACGACATCG..... 1040
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23  tProAtGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
   : : : : : : : : : : : : : : : : : : : : : : : :
1041 .....CTGGACACAGGCTCGGTGGCGGATCCA 1068
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40  subLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAsp 56
   : : : : : : : : : : : : : : : : : : : : : : : :
1069 ATTTGTAC.....GGCTGCAGCTCGGCGCAGCATCGCGTGGAC 1106
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57  AspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeuIle 73
   : : : : : : : : : : : : : : : : : : : : : : : :
1107 GGAGGAGCAGAGTGGCGGTCAACTCTTCGGCAGTGGCAGCG..... 1150
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73  yGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAla 90
   : : : : : : : : : : : : : : : : : : : : : : : :
1151 AGCAGCGGTCTACGACGCGCAACATGACTACTACTAC..... 1190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90  hrAlaProAsnMetPheAsnValAlaAspValLeuGlyValTyrSerPro 106
   : : : : : : : : : : : : : : : : : : : : : : : :
1191 .....AACAGCATGCAGCAGTACAGCGCG 1214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107  HisProTyrGluGlnGlnValSerAlaLeuGlyIleProTyr..... 121
   : : : : : : : : : : : : : : : : : : : : : : : :
1215 CCGCCCTTCTACTCGGATAC.....GGAACCTCTTATCGCGG 1252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 ..... 121
123  GCGCTGCTACTTGAAGCCAGCTATGCGCCAGCGGCAACACACTCG 1352
   : : : : : : : : : : : : : : : : : : : : : : : :
1303 CGGCTGCTACTTGAAGCCAGCTATGCGCCAGCGGCAACACACTCG 1352
   : : : : : : : : : : : : : : : : : : : : : : : :
123  GlnIleTyr..... 125
   : : : : : : : : : : : : : : : : : : : : : : : :
1353 CAGCTGTACAGCAGTCCGTACGCGGCTACACACTTCGCGCAGCAGA 1402

```

```

126 .....GlyTyrArgValAsnPheGly.ValIleAspGluArg... 138
   : : : : : : : : : : : : : : : : : : : : : : : :
1403 CTACGGCGGTACTACACAGCAGTACGCGCAGCTTACAGTCGGGCCA 1452
   : : : : : : : : : : : : : : : : : : : : : : : :
139 ...LeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeu 154
   : : : : : : : : : : : : : : : : : : : : : : : :
1453 ACTACTACCCGTATGGGTCAGCTCGCCAGCTCGAGTGGAGTCAATGA 1502
   : : : : : : : : : : : : : : : : : : : : : : : :
154 nIleAla.....ProAlaGlnAspGlyTyrArgGluAla.Gly 166
   : : : : : : : : : : : : : : : : : : : : : : : :
1503 CATGCTTCATGTGGCGGCTCTCGTAATCTCCGAGAGTCCACGGA 1552
   : : : : : : : : : : : : : : : : : : : : : : : :
167 PhePro.....ProAspHisGlnAlaTyrArg..... 175
   : : : : : : : : : : : : : : : : : : : : : : : :
1553 CACCACATCGACAGCCGCGTGCACAGACACCACACTCGCCGCACTCC 1602
   : : : : : : : : : : : : : : : : : : : : : : : :
176 .....GluGluProTyrPheHisAlaProGlnGly.... 186
   : : : : : : : : : : : : : : : : : : : : : : : :
1603 CGCTCCGATCTCGCCGAGCAGTGGCTCCGCGCATGGCCGCTGGCAAT 1652
   : : : : : : : : : : : : : : : : : : : : : : : :
187 .....CysGlyAsnSerSerArgThrIleThrGlyAspThr 198
   : : : : : : : : : : : : : : : : : : : : : : : :
1653 GTGTCCGCGCAGCTGGCGGCTGCTCTCACTCGACGCGAGCGACAG 1702
   : : : : : : : : : : : : : : : : : : : : : : : :
199 Cys 199
   |||
1703 TGT 1705

```

```

seq_name: /cgn2_6/ptodata/2/1na/PC/TUS.COMB.seq:PCT-US96-05611A-28

```

```

seq_documentation_block:

```

```

; Sequence 28, Application PC/TUS9605611A
; GENERAL INFORMATION:

```

```

; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Pelitrey, Clara M.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: Modified Myelin Protein Molecules
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage

```

```

; COMPUTER: Macintosh Centris 610

```

```

; OPERATING SYSTEM: System 7

```

```

; SOFTWARE: Microsoft Word 6.0.1

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05611A

```

```

; FILING DATE: 02-MAY-1995

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/431,644

```

```

; FILING DATE: May 2, 1995

```

```

; APPLICATION NUMBER: 08/431,648

```

```

; FILING DATE: May 2, 1995

```

```

; APPLICATION NUMBER: 08/482,114

```

```

; FILING DATE: June 7, 1995

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-129
; TELECOMMUNICATION INFORMATION:

```


TELEPHONE: (203) 255 1400
 TELEFAX: (203) 254 1101
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1476 base pairs
 TYPE: Nucleic acid
 STRANDEDNESS: Double
 TOPOLOGY: Linear
 MOLECULE TYPE: Other nucleic acid
 DESCRIPTION: MMOGp4 chimera
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 PCT-US96-05611A-28

alignment_scores:
 Quality: 80.50 Length: 319
 Ratio: 0.624 Gaps: 16
 Percent Similarity: 40.439 Percent Identity: 20.063

alignment_block:
 US-09-528-682-1 x PCT-US96-05611A-28 ..

Align seg 1/1 to: PCT-US96-05611A-28 from: 1 to: 1476

```

5 LeuTYrArgAlaAspSerArpProAspGluIleLysArgSerGly 21
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
247 ATGTACAAA...GACTCCACACACCGGCTCGTACCGCCATATAGGCTC 293
21 yLeuMetProArgGlyHisAsnGluTyrPheAsp..... 32
   :|||:|||||:|||||:|||||:|||||:|||||:
294 CCGCGCGGAGAAATCCACGGCGGTACCCAGATGAAAACCGGTGTCG 343
32 ..... 32
344 ACTTCTTCAAAAATTTGTGACCCCGCTACCCCGCGCTCTCAGGC 393
33 ArgGlyThrGlnMetAsnIleAsnLeuTyrAsp...HisAlaArgGly 48
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
394 AANAGCCGCGGCTGCTGCTGAGCCGTTTCAGCTGGGCGCGGAGGCCA 443
48 rGlnThrGlyPheValArgTyrAspAspGlyTyrValSerThrSerLeu 65
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
444 GGTCCGCGGCTC.....GGTACGCGGCGCGCTGCTCGC 478
65 eTLeuArgSerAlaHis..... 70
   :::::|||||:|||||:|||||:|||||:|||||:
479 ACTATTAATCTGCTCAAAAGCCTTCAAGCGGTGATGCCAGGCTACC 528
71 .....LeuAlaGlyGlnSerIleLeuSerGlyTyr 81
   :::::|||||:|||||:|||||:|||||:|||||:
529 TTGTCCAAAATTTTCAAACTGGGCGCGGTGATAGCCGTTTGGCTGCC 578
81 rThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnVal 98
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
579 G.....ATGCTAGAGCTCCCGGCGCAGTTTACAGTG 609
98 snAspValLeuGlyValTyrSerProHisProTyrGlnGlnGluValSer 114
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
610 .....ATAGGACCAAGACACCTATCCGGCTGTGTCGGG 645
115 AlaLeuGlyGlyLeuPro.....Tyr 122
   :::::|||||:|||||:|||||:|||||:|||||:
646 GATGAAGTGAATGCCATGTCCATATCTCTGGAGAGAACCTACAGG 695
122 rGlnIleTyrGlyTyrPyrArgValAsnPheGlyValIleAspGluArg 139
   :::::|||||:|||||:|||||:|||||:|||||:|||||:
696 CATGAGAGGGGTGCTACGCGCCCTTCTTACGGGTG...GTTCATC 742
139 euHisArgAsn.....ArgIleTyrArg 146
   :::::|||||:|||||:|||||:|||||:|||||:
743 TCTACAGAATATGCACAAGACCAAGATGAGACGACCTGAATATCGG 792
  
```

```

147 AspArg..... 148
   |||
793 GCGCGACAGAGCTGCTGAAGATGCTATTGGTGAGGGAAGTGCCTCT 842
149 .TyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeu 165
   |||||:|||||:|||||:|||||:|||||:|||||:
843 CAGGATCCGGAATGTAAAGTTCTCAGAT...GAAGAGGTTTCACCTGC 888
165 laGlyPheProProAspHisGln..... 172
   |||:|||||:|||||:|||||:|||||:|||||:
889 .....TTCTTCGAGATCATTTCTTACCAAGAGAGGACCAATGAATTG 933
173 .....AlaTyrArgGluGluProTyrPheHisAla 183
   |||:|||||:|||||:|||||:|||||:|||||:
934 AAAGTAGAAGATCCCTTCTACTGTGCTGAGAGATCCGGA...CATACAC 980
183 aProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 200
   | |||:|||||:|||||:|||||:|||||:|||||:
981 CCTCACTGGCAGACAGAAAGCTAATGAGACCTATTCTCCAAAACCTACC 1030
200 snGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlu..... 213
   :::::|||||:|||||:|||||:|||||:|||||:
1031 AAGACTATGATATCTCATCAATGATCATCTTCACAGTATGCTGAG 1080
214 .....TyrGlnSerIleValLysArgGlnIlePheSerAspTyrGln 228
   |||:|||||:|||||:|||||:|||||:|||||:
1081 GCGTTTACACACACCGCGCAGTACAGCAGATCTTGGGACTACAGAC 1130
228 rGluVal 230
   : |||
1131 CACCATC 1137
  
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-449-045C-23

```

seq_documentation_block:
: Sequence 23, Application US/08449045C
: Patent No. 5770203
: GENERAL INFORMATION:
: APPLICANT: Burnette, Neal W.
: TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
: TITLE OF INVENTION: SUBUNIT ANALOGS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Amgen Inc.
: STREET: 1840 De Havilland Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320-1789
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/449,045C
: FILING DATE: 24-MAY-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/694,733
: FILING DATE: 02-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/271,222
: FILING DATE: 06-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Mazza, Richard J.
: REGISTRATION NUMBER: 27,657
: REFERENCE/DOCKET NUMBER: A-196C
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 50 base pairs
: TYPE: nucleic acid
  
```


STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-449-045C-23

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-23 ..

Align seg 1/1 to: US-08-449-045C-23 from: 1 to: 50

179 TrrlEhISHisAlProGInGlyCysGlyAsnSerArgThr 193
|||||
3 TGGATTCAATCATGCACCGCGGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-24

seq_documentation_block:
; Sequence 24, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449, 045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-449-045C-24

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:

US-09-528-682-1 x US-08-449-045C-24 ..
Align seg 1/1 to: US-08-449-045C-24 from: 1 to: 50

179 TrrlEhISHisAlProGInGlyCysGlyAsnSerArgThr 193
|||||
3 TGGATTCAATCATGCACCGCGGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-35

seq_documentation_block:
; Sequence 35, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435, 605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
; US-08-435-605A-35

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-35 ..

Align seg 1/1 to: US-08-435-605A-35 from: 1 to: 50

179 TrrlEhISHisAlProGInGlyCysGlyAsnSerArgThr 193
|||||
3 TGGATTCAATCATGCACCGCGGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-36

seq_documentation_block:
; Sequence 36, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN


```

; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/435,605A
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: desc = "oligonucleotide"
; DESCRIPTION: /desc = "oligonucleotide"
US-08-435-605A-36

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-36 ..
Align seg 1/1 to: US-08-435-605A-36 from: 1 to: 50

179 TrrlshHsAlaProGlnGlycGlyAsnSerSeraArgThr 193
|||||
3 TGGATTTCATCATGCGACGCGAGGGTTGGGAGATGCTCCAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-128-450-25

seq_documentation_block:
; Sequence 25, Application US/09128450
; Patent No. 6211149
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Priola, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128,450
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Hamster sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (60)..(824)
US-09-128-450-25

```

```

alignment_scores:
Quality: 80.00 Length: 124
Ratio: 1.176 Gaps: 6
Percent Similarity: 54.839 Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-09-128-450-25 ..
Align seg 1/1 to: US-09-128-450-25 from: 1 to: 1000

114 SerAlaLeuGlyGlyIleProTySerGlnIleTyArgTrpTyArgVa 130
:::|||||
417 GCCCTGGTGGGGGCGCTT.....GGTGGCTACAGTCT 448
130 LAsnPhGlyValIleAspGluArgLeuHisArgAsnArgGluTyArgA 147
:::
449 GGGGAGTGGCCATGACAGCGCCCATGATGATTTGGCAATGACTGGGAGG 498
147 SPArgTyTrpArg...AsnLeuAsnIleAlaProAlaGluAspGlyTy 162
|||||
499 ACCGCTACTACCGTAAGAAACATGAACCGCTACCTTAACCAAGTATATAC 548
163 ArgLeuAlaGlyPheProAspHisGlnAlaTrpArgGluGluProTr 179
549 CGG.....CCAGTGGACCATGACAAACCAACGAGAACACTT 583
179 rPlshHsAlaProGlnGlycGlyAsn.....SerSeraArgT 193
|||
584 TGTGCAC.....GATTGTGTACATCATCACCATCAAGACAGACA 621
193 hrIleThnGlyAspThrCysAsnGlnGluTrpGlnAsnLeuSerThrIle 209
|||
622 CAGTCACCAACCAACCAACGAGGGAG.....AACTTACGAGAGAC 662
210 TyrLeuArgGluTyArgIleSerLeuValIleArgGlnIlePheSerAspTy 226
:::
663 GACATCAAGATATGAGACCGCGTGTGGAGACATGTGTACACCACTA 712
226 rGlnSerGluValAspIleTy 233
|||
713 TCAGAGAGAGTCCAGCGCTTAC 734

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-823-494-25

seq_documentation_block:
; Sequence 25, Application US/09823494
; Patent No. 6355610
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Priola, Joelle
; APPLICANT: Priola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/823,494
; PRIOR FILING DATE: 2001-03-30
; CURRENT APPLICATION NUMBER: 09/128,450
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Hamster sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (60)..(824)
US-09-823-494-25

alignment_scores:

```


Quality: 80.00 Length: 124
Ratio: 1.176 Gaps: 6
Percent Similarity: 54.839 Percent Identity: 24.194

alignment_block:

US-09-528-682-1 x US-09-823-494-25 ..

Align seg 1/1 to: US-09-823-494-25 from: 1 to: 1000

```

114 SerAlaLeuGlyLLeProtyrSerGlnIleTyrglyTPTyArgVa 130
      :::::::::::::::::::: ||| ||| ::
417 GCCGTGGTGGGGCCCTT.....GGTGGCTACATGCT 448
130 LAsnPhgIyValIleAspGluArgLeuHisArgAsnArgGluTyrgVa 147
      ::::: ::::: ::::: ::::: ::::: :::::
449 GGGGAGTGCATGAGCAGAGCCCATGATCATTTTGGCAATGATGCGGAGG 498
147 SPATGTYrTyArg...AsnLeuAsnIleAlaProAlaGluAspGlyTy 162
      ||||| ||||| ||||| ||||| ||||| |||||
499 ACCGCTACTACCGTGAACATGAAACCCCTACCCCTACCAAGTGATTTAC 548
163 ArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGluGluPro 179
      ||| ||| ||| ::::::::::: ::
549 CGG.....CCAGTGGACCACTACCAACAACCAAGCAACTT 583
179 PLeuHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
      ::::: ||| ||| ::::: ::::: :::::
584 TGTGCAC.....GATTGTGTCAACATCCATCAAGCAGCACA 621
193 hTleThGlyAspThrcysAsnGluGluThrGlnAsnLeuSerThrIle 209
      ||||| ||| ||| ::::: |||||
622 CAGTCCACCAACCACCAAGGGGGAG.....AACTTCACGAGAGACC 662
210 TyrLeuArgGluTyrglnSerLysValLysArgGlnIlePheSerAsp 226
      ::::: ::::: ||| ::::: :::::
663 GACATCAAGATATATGAGCGCGTGGTGGAGCAGATGTATACCAACCACTA 712
713 TCAGAAAGAGTCCACGAGCCTAC 734
226 rGlnSerGluValAspIleTy 233
      ||||| ||| ::::: |||
713 TCAGAAAGAGTCCACGAGCCTAC 734

```

seq_name: /cgn2_6/ptodata2/lna/6a.COMB.seq:US-08-728-603-14

seq_documentation_block:

Sequence 14, Application US/08728603
Patent No. 6093806

GENERAL INFORMATION:

APPLICANT: Cesarian, Ethel
APPLICANT: Knowles, Daniel M.
TITLE OF INVENTION: PROTEINS OF KAPOSI'S SARCOMA ASSOCIATED
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA

ZIP: 14603
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,603
FILING DATE: 10-OCT-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BRAMAN, SUSAN J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 19603/720

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ. ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4332 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-728-603-14

alignment_scores:

Quality: 80.00 Length: 263
Ratio: 0.630 Gaps: 18
Percent Similarity: 48.289 Percent Identity: 23.194

alignment_block:

US-09-528-682-1 x US-08-728-603-14/rev ..

Align seg 1/1 to reverse of: US-08-728-603-14 from: 1 to: 4332

```

14 AspLleuLysArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTy 30
      ||| ::::: ||| ||| ||||| ||||| :::::
3934 GATTGACTGGAAACCCATGGC.....CCACGGGACATGACAG.. 3896
30 rPheAspArgGlyThrGlnMetAsnIleAsnLeuTyArgHisAlaArg 47
      ||| ||| ::::: ||||| ||||| ||||| |||
3895 .....GGGTGTTAGGTGTGTGTGGAACATCTGCCATCCACGCA 3856
47 LyrThrGlnThrGlyPheValArgTyraAspAspGlyTyrgValSerThr 63
      ::::: ||| ||||| ||||| ||||| |||||
3855 TGAACATCAAGATAGGTC...TATGATCAGGCGACGTCGCCGCCGTC 3809
64 LeuSerLeuArgSerAlaHis.....LeuAlaGlyGln 75
      ::::: ||||| ||||| ||||| ||||| |||||
3808 ACTGATGTATAGTCCGCCACCCGTGAGTTGCCGTGGGGTTCTGGGAT 3759
71 .....LeuAlaGlyGln 75
3758 AGTGTCTGGCGGCGAGGTCATCCCGGCATTTCCATGTAGTACG 3709
75 erLleuSerGlyTySerThrTyrrIle.....TyrrAlleAla 89
      ::||| ::::: ||| ||| ::::: :::::
3708 GTATATCTCGCTCGCTGCTCAGTATGATCTACGAGGCGTCGTCGTGA 3659
90 ThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTySerP 106
      ||||| ::::: ||||| ||||| ||||| |||||
3658 CCGGACCCCGGAGTACTCTCCCTGGGCCAGCT...GGGACAGACCGCTC 3612
106 O.....HisProTyrgluGlnGluValSerAlaLeu.GlyGlyIlePro 120
      ||| ||| ::::: ||||| ||||| ||||| |||||
3611 CCCGCAACACTCGGAGGAAAAAGCTTATGTTCTGAGGATCTGATG 3562
121 TySerGlnIleTyrglyTyrgValAsnPhgIyValIleAspG 137
      ::||| ||||| ||||| ||||| ||||| |||||
3561 TTATGCCAG.....TGGCTCATACAGCTTGACACAGTTGGTCTC 3521
137 uArgLeuHisArgAsnArg.....GluTyArgAspArgT 149
      ::||| ::::: ||| |||
3520 CAGGTTTACCGCCCGAGCGCTGGGTGTGGTCCCTCA.....G 3480
149 YrTyArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrgValLeuAl 165
      ||| ::::: ||||| ||||| ||||| ||||| |||||
3479 TGTATGTGAGGATTCGACCGGCGCCCTACACCCAGGGCAGCAGCAGC 3430
165 aglyPheProProAspHisGlnAlaTyrArgGlu..... 176
      ||||| ||||| ||||| ||||| ||||| |||||
3429 TGAAGCCACCTCGCCACAGAGATGAGATGTGTGCGGTCTGTTAG 3380
177 .....GluProTyrIleHisHis.....AlaPro.G 185
      ::||| ||||| ||||| ||||| ||||| |||||
3379 AAACCTGTACAGGTGAGGACACAGGTAGGTCGTATACACAGGCGCAGA 3330

```



```

30 rphesapargglyrhgrlnmetanllasneulyrasphlsalarg 47
25328 .....GGGGTTAGAGTCTGGTGGACATGGCTGCCATGCCACGGA 25367
47 lythrglnthnglyrphavalargtyraspaspgltyvalserthser 63
25368 TGGAACTCATCAGAGATGGGCT...TATGATCAGGGCAGCGCTGTCGCCGCTC 25414
64 Leuserleuabrgseralanhis..... 70
25415 ACTGAGATGAAGTCCGGCCACCGTGGAGTTCGCTGTGGGTTTCTGGAGT 25464
71 .....leualaglylgs 75
25465 AGTGTCTGGTGGCAGGCTTCATCCGGCGCATTTCCATGGTAGGTGAGG 25514
75 erlleuserglytyrserthrtyrtyrle.....tyrvalilleala 89
25515 GTTATCTCGGCTGCTGTCTGATGATGACAGAGGCGCTGCTGTCGCTA 25564
90 ThAlaProasnmctrphesnaValasnaPyalaLeuGlyValtyrSerPr 106
25565 CCGGACCCCGAGGATCTCTCCCTGGGCCAGCT...GGGCAGACCCGCTC 25611
106 o.....HisProtyrgrnglnlvalserlaleu glyglyllepro 120
25612 CCGCGAACCCTCGGAGGAAACGCTTGAATGTTCTGAGAGGATCTGTATG 25661
121 tyrserglnlleyrgltyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 137
25662 TTTAAGCCAG.....TGCGTGCATACACCTTGGACACGTTGGTCTC 25702
137 uargleuHisargAsnarg.....GluTyraArgAspArgT 149
25703 CAGGTTTACCGCGCCAGCGCTGGGGTGTGGGCTCGTAC.....G 25743
149 ylyrtyrarg AsnleuAsnlllealprolagluaspgltyrtyrtyrleu 165
25744 TGTATGTGAGGATTCGACCGGCCCACTTACACCCAGGGCCACACGACG 25793
165 aglyrpheroProasphlslnalatrprarglu..... 176
25794 TGGAGCCCACTCGCCACACAGATGGAGATGTGTGGGTCTGTTAG 25843
177 .....GluProTPrllleHshis.....Alapro.Gl 185
25844 AAACCTGTGACGGTGGAGGACAGGTAGGGTGCATTACAGCGCCAGGA 25893
185 n.....GlyysgilyanserSerArghrllleThnrglyAspDhnc 199
25894 CCGATCCCTGGCGCTGGCGTAGCTGGCGCTGGCAGCGCTTTCGACACAT 25943
199 ysaSn.....GluGluThrglnAsnleuSerThrllleTyrlleu... 211
25944 GTATATCAGACAGAGAACCCGACAGAGACTGCTCGTTTAAGCTCTTC 25993
212 ArgGluTyrglnserlysvallysarglnlle 222
25994 CACAGTCAACGCTGCACACTCAAAAGCCCGCTGT 26026
seq_name:/cgn2_6/plodata/2/lna/6h_comb_seq.us-08-757-669A-20
seq_documentation_block:
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS

```


Align seg 1/1 to: US-09-230-371A-20 from: 1 to: 32207


```

628 AGTTCGAATACATATTCACGGCCACATCCAGTCATGATATATCTCAGT 677
115 a.....LeuGlyIleProTySerGlnIleTyGlyT 127
      :|||||:
678 ATTCAAGAATTCGAACGTAGTGTGCTTCATGAAACAACGTAGT 727
      :|||||:
127 rPTyArGyAlaSnPheGlyValIleAspGluArgLeuHisArgAsnArg 143
      :|||||:
728 ACAGTCAGTAATGAAGCAATATTGAAAGAAATTCAAAATAGTAGCGGT 777
      :|||||:
144 GluTyArGAspArgTyTyArg.....AsnLeuAsnIle..... 155
      :|||||:
778 TCCTATACGTCAAAAATTTTAACTTACTACACCTTGCTGCTATCTCTTA 827
      :|||||:
156 ....AlaProAlaGluAspGlyTyTyArgLeuAlaGlyPheProProAspH 171
      :|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
      :|||||:
171 ts...GlnAlaTrpArgGluGluProTrpIleHisAlaProGlnGly 186
      :|||||:
878 ACACAGAAAAGTGGAGAAAATTTATCTTGATCTATCATTCATGCTCT 927
      :|||||:
187 CysGlyAsnSerSerArgTrpIleThrGlyAspTrpCysAsnGluGluTh 203
      :|||||:
928 GATGGTCGGTCTTCGATCCACTTTTTCATGAT...TTAAGAGAGCAAT 974
      :|||||:
203 rGlnAsnLeuSerThr.....IleTyTrpLeuArg 213
      :|||||:
975 AATATATATTAATAACACACCAAAAATTAAGATTAATTCATTCACAGTAC 1024
      :|||||:
213 lUTyGlnSerValLysValArgGlnIlePheSerAspTyTrpGlnSerGlu 229
      :|||||:
1025 AGGAGATTCACCAATTATTCAGGAACTTCAGAACCGATCGAAAAGGTG 1074
      :|||||:
230 ValAsp..... 231
      :|||||:
1075 ATAGACTTATGACACCGTACTGTTTATTCGAACTCACTTCTTCGGG 1124
      :|||||:
232 ....IleTyAsnArgIleArg 237
      :|||||:
1125 TTTCATCTACAAATCATTTGAGA 1146
      :|||||:
seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-461-599-16
seq_documentation_block:
: Sequence 16, Application US/08461599
: Patent No. 5658777
: GENERAL INFORMATION:
: APPLICANT: FUJII, Toshio
: APPLICANT: IWAMATSU, Akihiko
: APPLICANT: KOSHIMOTO, Hiroyuki
: APPLICANT: MINETOKI, Toshitaka
: APPLICANT: BOGAKI, Takayuki
: APPLICANT: NAGASAWA, Naoshi
: TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,599
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/077,939
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 62997/1993
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 184328/1992
FILING DATE: 18-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/105 KYPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1974 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 346..1923
US-08-461-599-16

alignment_scores:
Quality: 78.00 Length: 208
Ratio: 0.743 Gaps: 10
Percent Similarity: 50.481 Percent Identity: 22.115

alignment_block:
US-09-528-682-1 x US-08-461-599-16 ..
Align seg 1/1 to: US-08-461-599-16 from: 1 to: 1974

68 SerAlaHisLeuAlaGlyInSerIleLeuSerGlyTySerThrTyTy 84
|||||:
528 TCAGCTCACATTTAGCTTGGAGGAAATCTGCGTGAATAATCCAACTTT 577
|||||:
84 rIleTyValIleAlaThrAla...ProAsnMetPheAsnValAsnAsp 99
|||||:
578 TACATATTGTTCTTACACACAGATGCCCAATCATCAAAATTTTATCC 627
|||||:
100 ValLeuGlyValTySer...ProHisProTyGluGlnGluValSerAl 115
|||||:
628 AGTTCGAATACATTCACGGCCACATCCAGTCATGATATATCTCAGT 677
|||||:
115 a.....LeuGlyIleProTySerGlnIleTyGlyT 127
      :|||||:
678 ATTCAAGAATTCGAACGTAGTGTGCTTCATGAAACAACGTAGT 727
      :|||||:
127 rPTyArGyAlaSnPheGlyValIleAspGluArgLeuHisArgAsnArg 143
      :|||||:
728 ACAGTCAGTAATGAAGCAATATTGAAAGAAATTCAAAATAGTAGCGGT 777
      :|||||:
144 GluTyArGAspArgTyTyArg.....AsnLeuAsnIle..... 155
      :|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
      :|||||:
778 TCCTATACGTCAAAAATTTTAACTTACTACACCTTGCTGCTATCTCTTA 827
      :|||||:
156 ....AlaProAlaGluAspGlyTyTyArgLeuAlaGlyPheProProAspH 171
      :|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
      :|||||:
171 ts...GlnAlaTrpArgGluGluProTrpIleHisAlaProGlnGly 186
      :|||||:
878 ACACAGAAAAGTGGAGAAAATTTATCTTGATCTATCATTCATGCTCT 927
      :|||||:
187 CysGlyAsnSerSerArgTrpIleThrGlyAspTrpCysAsnGluGluTh 203
      :|||||:
928 GATGGTCGGTCTTCGATCCACTTTTTCATGAT...TTAAGAGAGCAAT 974
      :|||||:

```



```

US-08-461-621-16

alignment_scores:
    Quality: 78.00      Length: 208
    Ratio: 0.743      Gaps: 10
    Percent Similarity: 50.481      Percent Identity: 22.115

alignment block:
US-09-528-682-1 x US-08-461-621-16  ..

Align seg 1/1  to: US-08-461-621-16  from: 1  to: 1974

68  SerAlaHisLeuAlaGlyInserIleLeuSerGlyTySerThrTyTy  84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528  TCAGCTCACATTAGCTTGGGGAACCTGCCGTAATAATCCAACTCTT  577
84  rIleTyValIleAlaThrAla...ProAsnMetPheAsnValAsnsp  99
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
578  TACATATTGTTCTTACCAACAAGATGGCCAAATCATGAAATATTATTCG  627
100  ValLeuGlyValTySer...ProHisProTyGluGlnIleValSerAl  115
628  AGTTCGGAATACTATTTCACGGCCACATCCGTCGATCATATTATCTCAG  677
115  a.....LeuGlyGlyIleProTySerGlnIleTyGlyTy  127
   :|||||:|||||:|||||:|||||:|||||:|||||:
678  ATTACAAAGATTGAACCTGAGTGTGTGTTCTTCATGACCAACCTGAGT  727
127  rPTyTyTyValAsnPheGlyValIleAspGluTyGleuHisArgAsnArg  143
   :|||:|||||:|||||:|||||:|||||:|||||:
728  ACGATGCAGTATGAAGCAAAATATTGAAGAATTTCAAAATAGTAAGGCT  777
144  GluTyTyArgAspArgTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy  155
   :|||||:|||||:|||||:|||||:|||||:|||||:
778  TCTTATACGTCGCAAAATTTTAAACTTACTACACACTTGTGACTATTCCTTA  827
156  ...AlaProAlaGluAspGlyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy  171
   :|||||:|||||:|||||:|||||:|||||:|||||:
828  CTTTGGACCAACGAGCCGATGTCGGCGCTTAATTTGCTTCCCAAGAGAGC  877
171  is...GlnAlaThrPArgGluGluProThrIleHisAlaIleProGlnGly  186
   :|||:|||||:|||||:|||||:|||||:|||||:
878  ACGACAGAAAGTGGAGAAATTTATCTTTGTATGTTAAATCATTCATGATGCT  927
187  CysGlyIleAsnSerSerArgThrIleThrGlyAspThrCysAsnGluTyTh  203
   :|||||:|||||:|||||:|||||:|||||:|||||:
928  GATGGTCGGCTTCGATCCACTTTTTCATGAT...TTAAGAGAGCAAT  974
203  rGlnAsnLeuSerThr.....IleTyTyLeuArg  213
   :|||||:|||||:|||||:|||||:|||||:|||||:
975  AATATATTATTAAACGTCCACCAAAAAATTAGATTCAATTTTCAAGTACG  1024
213  IuTyGlnSerIleValIleAspGlnIlePheSerAspTyTyGlnSerGlu  229
   :|||:|||||:|||||:|||||:|||||:|||||:
1025  AGGAGATTAACCAATTATTATGAGAACTTCAGAAACGATCGAAAGGTG  1074
230  ValASP.....  231
   :|||||:|||||:|||||:|||||:|||||:|||||:
1075  ATAGACTTTAGACCACCGTACTGTTTATTCGGAAGTCACTTCTTTCGGG  1124
232  ....IleTyTyAsnArgIleArg  237
   :|||||:|||||:|||||:|||||:|||||:|||||:
1125  TTTCATCTACATCATCTTTTGAGA  1146

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-334-16
seq_documentation_block:
; Sequence 16, Application US/08465334
; Patent No. 5728412
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IWAMATSU, Akihito

```



```

APPLICANT: YOSHIMOTO, HIROYUKI
APPLICANT: MINETOKI, TOSHITAKA
APPLICANT: BOGAKI, TAKAYUKI
APPLICANT: NAGASAWA, NAOSHI
TITLE OF INVENTION: ALCOHOL ACETYLTANSFERASE GENES AND USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,334
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,939
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 62997/1993
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 184328/1992
FILING DATE: 18-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/101 KYPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1974 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 346..1923
US-08-465-334-16

alignment_scores:
Quality: 78.00 Length: 208
Ratio: 0.743 Gaps: 10
Percent Similarity: 50.481 Percent Identity: 22.115

alignment_block:
US-09-528-682-1 x US-08-465-334-16 ..
Align seg 1/1 to: US-08-465-334-16 from: 1 to: 1974

68 SetAlHIsleuAlaGlnSerIleLeuSerGlyTyrSerThrTyrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 TCAGCTCACATTTAGCTTTGAGGAAATCTGCGTGAATAATCCACACTCTTT 577
84 rIeTyrValIleAlaThrAla...ProAsnMetPheAsnValAsnAsp 99
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578 TACATATGTGTTCACCAACAAGATGCGCAATATCATGAAATTTATATCGC 627
100 ValIleuGlyValTyrSer...ProHisProTyrGluGlnGluValSerAl 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 AGTTCGAATACTATTCACGCGCACATCCAGTCAGATGATATATCTCAGT 677

```

```

115 a.....LeuGlyGlyIleProTyrSerGlnIleTyrGlyT 127
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
678 ATTACAGAATTGAACCTGAGTGTGCTTCTCAATGAAACAACCTGACT 727
127 rPTyrArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArg 143
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
728 ACAGTCAGATGATGAGCAACAATATTGAGAGAAATTCAAAATGATGAAGCT 777
144 GluTyrArgAspArgTyrTyrArg.....AsnLeuAsnIle..... 155
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 TCCATATCTGCAAAAATTTTAACTTACTACCACTTTGACTATCTCTTA 827
156 ...AlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsp 171
828 CTTTGGACACAGACGACGAGTGGCGCTAATTTGCTTCCAGAAAGCG 877
171 Is...GlnAlaTyrArgGluGluProTyrIleHisAlaProGlnGly 186
||:|||||:|||||:|||||:|||||:|||||:|||||:
878 ACACAGAAAGTGAGGAAATTTATCTTGTATCTATCATCTTCATGCTCT 927
187 CysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsnGluGlu 203
|||:|||||:|||||:|||||:|||||:|||||:|||||:
928 GATGTCGCGCTTCGATCCACTTTTTCATGAT...TTAAGACAGCAAT 974
203 rGlnAsnLeuSerThr.....IleTyrLeuArg 213
:|||||:|||||:|||||:|||||:|||||:|||||:
975 AATTAATATTAAACCTCCACCAAAAATTAATTAATTTCAATTTCAAGTAC 1024
213 IuTyrGlnSerIleValIleAspGlnIlePheSerAspTyrGlnSerGlu 229
||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 AGAGAGATTACCAATTTATGAGAAACTTCACAGACGATCGAAAGGCT 1074
230 ValAsp..... 231
:|||||:|||||:|||||:|||||:|||||:|||||:
1075 ATAGACTTTAGACACCGTACTGTTTATTCGAAAGTCACTCTTTCGGG 1124
232 ...IleTyrAsnArgIleArg 237
|||||:|||||:|||||:|||||:|||||:|||||:
1125 TTTCATCTTACATCATTTTGAGA 1146

seq_name: /cgn2_6/plodata/2/1na/6A_COMB.seq:US-08-990-470A-1

seq documentation block:
; Sequence 1, Application US/08990470A
; Patent No. 6123942
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: SO, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nassif, Xavier
; TITLE OF INVENTION: No. 6123942el Bacterial Hemoglobin Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDowell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,470A
; FILING DATE: 15-DEC-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6123942nan, Kevin E
; REGISTRATION NUMBER: 35,303

```



```

REFERENCE/DOCKET NUMBER: 94,784-K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3318 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 470..2845
US-08-990-470A-1

alignment_scores:
Quality: 77.50 Length: 369
Ratio: 0.601 Gaps: 15
Percent Similarity: 34.959 Percent Identity: 18.157

alignment_block:
US-09-528-682-1 x US-08-990-470A-1 ..

Align seg 1/1 to: US-08-990-470A-1 from: 1 to: 3318

7 ArgAlaSPSerArgProProAspGluIleLysArgSerGlyGlyLeu 23
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1601 CGTTGGACACCATCCG.....TTGCACCTCGGGGGGGCGACA 1641
23 tProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
::: ::::: ||||| ||||| ||||| ||||| ||||| |||||
1642 CCGCGCTGCGTTAAACTTCCGCCAGCGCGCGATTTGAAACCTCA 1691
40 snLeuTyrAspHis..... 44
||| |||||
1692 ACCCGACGATTATTACTTTCAGCGCGGTGTTGTCGAACCAACAGCAG 1741
45 .....AlaArgGlyThrGlnThrGlyPhe..... 52
||||| ||||| ||||| ||||| ||||| ||||| |||||
1742 ATCCAGCATCCGGTGAAACCACTACGCTTCTCAGTGTGTGACCA 1791
53 .....ValArgTyrAsp..... 56
||||| ||||| ||||| ||||| ||||| ||||| |||||
1792 AATTCAATGACGACGCTTCACTAGCCGCGCAGGATCCGTTACGATC 1841
56 ..... 56
1842 ATACCAAAATGACCGCTCAGGAATGAATGCCAGTGCATGCTTGAC 1891
57 .....AspGlyTyrValSer..... 61
1892 AAAACACCGCTCGACGCCACACTTATTAAGCGTGAGCGGTTTGTGCG 1941
61 tThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
::: ::::: ||||| ||||| ||||| ||||| ||||| |||||
1942 CTGGCGCGCGCACTGATGATGAGGCTTGCGGTGAGTACACATTTACT 1991
78 ergGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsn... 93
||||| ||||| ||||| ||||| ||||| ||||| |||||
1992 CCGGCTAC.....CGTGTCGCCCAATGCG 2014
94 .....MetPheAsnValAsnAspValLeuGlyValTyrSerProH 107
||||| ||||| ||||| ||||| ||||| ||||| |||||
2015 TCCGAAGTGTATTCTACTTACACACGCGTTGGGTATGTCGCCAA 2064
107 sPro..... 108
|||||
2065 TCCCAACCTGAAAGCCGACGCGACGACCAACCCCTCTCTGCAAG 2114
108 ..... 108

```

```

2115 GCCGACGAAAAAGTACTTGGATGCCAACTGATCAAAAGCAATFAC 2164
109 .....TyrGlnGlnGluValSerAlaLeuGlyIle... 119
||||| ||||| ||||| ||||| ||||| ||||| |||||
2165 CGCAATTCTGTCTGAGAGCAGAACTGACCAAGCGGCGCATGTCAAG 2214
120 .....ProTyrSerG 123
||||| ||||| ||||| ||||| ||||| ||||| |||||
2215 CTGTACTGACATGATTTACTACTAGGTATGTAGCAATCTTATTTCG 2264
123 InIleTyrGlyTyrTyrArgValAsnProGlyValIleAspGluArgLeu 139
::: ::::: ||||| ||||| ||||| ||||| ||||| |||||
2265 AAAAATCGAATGCGAGATGCAAAAT.....ATCGACAGGCCACAG 2305
140 HisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuValIle 156
||||| ||||| ||||| ||||| ||||| ||||| |||||
2306 ATCCGCGCTCGAGCTGACGGGCGCT.....CTGATGTGGA 2343
156 aProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGln 173
||||| ||||| ||||| ||||| ||||| ||||| |||||
2344 C.....AAAGTAGCGTCTTTGTTCT..... 2365
173 LatrParGluGluProTyrPheHisAlaProGlnGlyCysGlyAsn 189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2366 .....GAGGCGTGAAACTGTCGCTGCGTGGGTTATGCGAAA 2404
190 SerSerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsn 206
||||| ||||| ||||| ||||| ||||| ||||| |||||
2405 AGCAAA.....CTGTGCGCGACAAACAGCTGCTGTCACCAAGCGGT 2448
206 userThrIleTyrLeuArgGluTyrGlnSer..... 216
||||| ||||| ||||| ||||| ||||| ||||| |||||
2449 GAAAGTATTTCGCGTATGATGAACTCCGAGGAAAAAATGGGGCG 2498
217 .....LysValLysArgGlnIle 222
||||| ||||| ||||| ||||| ||||| ||||| |||||
2499 TGTTCTCCCGCTGACCTATCTGCGCGCAAAAAGTCAAAGACGCGCA 2548
223 PheSerAspTyrGlnSerGlu.....ValAspIleTyrAsnArgI 236
||||| ||||| ||||| ||||| ||||| ||||| |||||
2549 TACACGTTTATGAAACAAAGCGCTGGGTTAGCGCTTTGCAAAAAAG 2598
236 eArgAsp 238
|||||
2599 AAAAGAT 2605

seq_name: /cgn2_6/ptodata/2/lna/6B.COMB.seq:US-08-817-707-1

seq_documentation_block:
; Sequence 1, Application US/08817707
; Patent No. 6277382
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: So, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nessif, Xavier
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

```


94MetPheAsnValAsnAspValLeuGlyValTyrSerProH
|||:: ||| ||| :: |||::

; LENGTH: 4460

```

;  LENGTH: 4
;  TYPE: DNA

```


36 GIMetAsnIIleasLeuTYaspHisLa.....ArgGlyYthrGI 49
:::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
3153 AATTAAGAATAACTTCCTTCTAATAAACATTGTTAACCAAGATAGTCA 3202

; PRIOR FILING DATE: 1998-08-03
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 27
 ; LENGTH: 1322
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(865)
 ; US-09-823-494-27

alignment_scores:
 Quality: 77.00 Length: 124
 Ratio: 1.149 Gaps: 6
 Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:
 US-09-528-682-1 x US-09-823-494-27 ..

Align seg 1/1 to: US-09-823-494-27 from: 1 to: 1322

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrTyrGva 130
    ::::::::::::::::::::|||:::
455 GCAGTAGTGGGGGGCTT.....GGTGGCTACATGCT 486
130 LAsnPhgGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
    ::::::::::::::::::::|||:::
487 GGGGAGCGCCGTGTAGACAGGCCCATATTCATTGGCAACGACGTGGAG 536
147 SPATYTYTYTArg..AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
    |||||:::|||||:::
537 ACCGCTACTACCGTGAACATGTACCGCTACCTCAACCAAGTGTACTAC 586
163 ArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGluTyr 179
    |||:::|||||:::
587 AGG.....CCAGTGCATCAGTACAGCAACCAACAACACTT 621
179 pIleHisIleAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
    ::::::::::|||::|::|::|::|
622 CGTGCAC.....GACTCGTCATATATCATCATTCAAGCAGACCA 659
193 hrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    ||::|::|::|::|::|::|::|
660 CGGTACACACACACCAAGGGGAG.....AACTCACCAGACAC 700
210 TyrLeuArgGlyTyrGlnSerLysValLysArgGlnIlePheSerAspTyr 226
    ::::|::|::|::|::|::|::|
701 GATGTGAAGATGATGAGCGCGTGTGAGACGATGTGCTCACCCAGTA 750
226 rGlnSerGluValAspIleTyr 233
    |||::|::|::|::|::|
751 CCAGAGAGATGCCAGGCGCTAT 772
  
```

seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-326-670A-1

seq documentation block:

; Sequence 1, Application US/08326670A
 ; Patent No. 5698438
 ; GENERAL INFORMATION:
 ; APPLICANT: Stojiljkovic, Igor
 ; APPLICANT: So, Magdalene
 ; APPLICANT: Hwa, Vivian
 ; APPLICANT: Hefron, Fred
 ; APPLICANT: Nassif, Xavier
 ; TITLE OF INVENTION: A No. 5698438el Bacterial Hemoglobin Receptor Gene and Uses
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Allegrretti & Mitcoff, ltd.
 ; STREET: 10 South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: Illinois

; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/326.670A
 ; FILING DATE: 18 OCT 1994
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5698438nan, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 94,784
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 810-221-8317
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3318 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 470..2848
 ; US-08-326-670A-1

alignment_scores:
 Quality: 76.50 Length: 369
 Ratio: 0.593 Gaps: 15
 Percent Similarity: 34.959 Percent Identity: 18.157

alignment_block:
US-09-528-682-1 x US-08-326-670A-1 ..

Align seg 1/1 to: US-08-326-670A-1 from: 1 to: 3318

```

7 ArgAlaAspSerArgProProAspGluIleLysArgSerGlyLeu 23
    ||| |||||:::|||||
1601 CGTTTGACAGCCATCGC.....TTGCACCTGGGGGGGGGAGACA 1641
23 tProAlaGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
    ::::|::|::|::|::|::|
1642 CCGCGTGTGCTTAAACCTTTCGCGACGCCGCGTGATTTTGAACCTTA 1691
40 snLeuTyrAspHis.....
1692 ACCGCGACGATTTTACTTTCAGTACGCCGCGTGTGTTGAACCCACACAGT 1741
45 .....AlaArgGlyThrGlnThrGlyPhe.....
1742 ATCCAGCATCCGTTGAACCAACCACTACGAGTTTCACTGTCAGCA 1791
53 .....ValArgTyrAsp 56
1792 AATTCAATGAGACAGCGTGTTCAGTAGCCGCGCAGGATTCGTTAGCAGTC 1841
56 .....
1842 ATACCAAAATGACCGCTCAGGAATTGAATGCCAGATGTCATGCTTGTGAC 1891
57 .....AspGlyTyrValSer 61
1892 AAAACACCGCTGCAGCCACTTAAAGCTGGAGCGGCTTTGTGCG 1941
61 rThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleAsn 78
    ::::|::|::|::|::|::|
1942 CTTGGCGGCGCAACTGATCAGCGCTTGGCGTGTGCGTTACGACATTA 1991
  
```



```

78  erGIYrSerThrTYrIleYrValIleAlaThrAlaProAsn... 93
    |||||
1992  CCGGCTAC.....CGTCTCCCAATGCG 2014
94  .....MetPheAsnValAsnAspValIleuGIYrSerProH 107
    |||||
2015  TCGAAGTGTATTCTACTACAAACGAGGTTCGGTAAATGGCGCCAA 2064
107  sPro.....
    |||
2065  TCCCAACCTGAAGCCGAGCGACAGCACACCACCCCTCTCTGCAAG 2114
108  .....
2115  GCGGACGCAAAAAGTACTTTGGATCCCAACCTGATCAAGCAATAC 2164
109  .....TyrGIuGIuValSerAlaIleuGIYrIle... 119
    |||||
2165  CGCAATTTCTGTCTGAAGAGCAAGACTGACCAAGCGCGCATGTGAG 2214
120  .....ProTyrSerG 123
2215  CTGTACTCAGATGAAATTACTACTAGCGGTATGTAGCAATCTTATCCG 2264
123  InIleTYrGIYrTYrTYrValAsnPhelGIYrValIleAspGIuArgLeu 139
    |||||
2265  AAAAAGCTGAATGGCAGATGCAGAAAT.....ATCGAACAGCGCAGA 2305
140  HisArgAsnArgGIuTYrArgAspArgTYrTYrArgAsnLeuAsnIleAl 156
    |||||
2306  ATCCGCGGTATTCAGAGCTGACGCGCGCT.....CTGATGTGGA 2343
156  APpOAlaGIuAspGIYrTYrArgIleuAlaGIYrPheProProAspHisGIuAl 173
    |||||
2344  C.....AAAGTACGCTCTTTGTTCCT..... 2365
173  IaTYrArgGIuGIuProTYrIleHisHisAlaProGIuGIYrGIYrAsn 189
    |||||
2366  .....GAGGCTGGAAAGCTGTCGCGCTCGCGGTATGCGAA 2404
190  SerSerArgThrIleThrGIYrAspThrCysAsnGIuGIuThrGIuAsnLe 206
    |||||
2405  AGCAAA.....CTGTGCGCGGACAAACAGCCTGTCTCCACCCAGCCGT 2448
206  UserThrIleTYrIleuArgGIuTYrGIuSer..... 216
    |||||
2449  GAAAGTGTATGCGGTATCGACTATGAAGTCCGAGCAAAAATGGGCG 2498
217  .....LysValIysArgGIuIle 222
2499  TGTTCCTCCCGCTGACCTATCTGGCGCGGAAAGGTCAAAGACGCCAA 2548
223  PheSerAspTYrGIuSerGIu.....ValAspIleTYrAsnArgI 236
    |||||
2549  TACACCGTTTATGAACAAGAGCGGTGGGTACGCTTTGCAGAAAAAGCT 2598
236  eArgAsp 238
    |||||
2599  AAAAGAT 2605
seq_name: /sgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-482-918-47
seq_documentation_block:
; Sequence 47, Application US/08482918
; Patent No. 6207417
; GENERAL INFORMATION:
; APPLICANT: zsebo, Krisztina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Martln, Francis H.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gershtein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482, 918
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/33005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 5864 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: join(565..579, 1684..1797, 2693..2755, 3351..3521,
LOCATION: 3932..4088, 4314..4397, 4778..4887, 5208..5275,
LOCATION: 5677..5713)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: join(1744..1797, 2693..2755, 3351..3521, 3932
LOCATION: 4088, 4314..4397, 4778..4887, 5208..5275, 5677
LOCATION: ..5713)
US-08-482-918-47
alignment_scores:
Quality: 76.50 Length: 214
Ratio: 0.850 Gaps: 12
Percent Similarity: 42.056 Percent Identity: 25.234
alignment_block:
US-09-528-682-1 x US-08-482-918-47 ..
Align seg 1/1 to: US-08-482-918-47 from: 1 to: 5864
11 ArgProProAspGIuIleIysArgSer..... 19
|||||
245 CGCCTCCGATTCCTCCCTCCGCGCCCTGCTTCCTTCGCCCTACCCC 294
20 .....GlyGIYLeuMetProArgGIYHisAsnGIuTYrP 31
|||||
295 GGCCTCGGAAGGAGAGAGCGGTGTCCGAGCAGCGCGGGGGA.... 340
31 heAspArgGIYThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGIY 47
|||||
341 .....CTGTATAAAGCGCGCGCGC 361
48 ThrGlnThrGIYrPheValArgTYrAspAspGIYrValSerThrSerIe 64
|||||
362 TCACACCGCGCTTCGCTCGC.....CGCCTCGGCGCGAGACT 399
64 u...SerLeuArgSerAlaHisLeuAlaGIYGIuSerIleuSerGIY 80
|||||
400 AGAAGCGCTGCGGAGACAGGAGACGTGAGAGAGCGCGCTCGCGGCT 449

```



```

132 heglyValIleaspluarleuHisArgAsnArgLutyrArgsparg 148
613 .....TCGAGAAAAATCGCCCGG 633
149 TyrTyrArgAsnLeuAsnIleAlaProAlaGluSplGlyTyrArgLeuAl 165
634 .....TGCCCCGGGAA.....GC 647
165 aGlyPheProAspHisGlnAlaTPrArgGluGluProTrpIleHisH 182
648 CGGCGCTCCCTGGAGCTTGCACGCTGGCGGTGCAGCGCTGTGCTGCCG 697
182 IsAlaProGlnGlyCysGlyAsnSerArgTyrIle 194
698 GTGACAGAAAGATCGGGGGAGCGCGCGTGTGTG 735

```

```

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-537-361E-5

```

```

seq_documentation_block:
; Sequence 5, Application US/08537361E
; Patent No. 6121037
;
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: So, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nasif, Xavier
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor
; TITLE OF INVENTION: Genes and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,361E
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6121037nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 94,784-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2376
;
; US-08-537-361E-5

```

alignment_scores:

```

Quality: 76.00 Length: 292
Ratio: 0.710 Gaps: 13
Percent Similarity: 36.644 Percent Identity: 19.521

```

alignment_block:

```

US-09-528-682-1 x US-08-537-361E-5

```

Align seg 1/1 to: US-08-537-361E-5 from: 1 to: 2379

```

42 TyrAspHis.....AlaArgGlyThrGlnThrGlyPheValArgT 44
1366 TAGCATCATACCAAAATAGCGCCTCAGAAATGAAATGCCAGTGTATGC 1415
45 .....AlaArgGlyThrGlnThrGlyPheValArgT 55
1416 TTGTGACAAAACACCGCTTCAGCAGCACTTATAAGCTGG..... 1458
55 yraspsplGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeu 71
1459 .....ACGGTTTGTGCTGGCTGGCGGCGCACTGATCAGCTTGCGC 1503
72 AlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValI 88
1504 GTCGGTTACGACATTAATCTTCCGCTAC..... 1530
88 eAlaThrAlaProAsn.....MetPheAsnValAsnAspValI 101
1531 ....CGTGTCCTCCATGCGTCCGAGATGTATTCTACTTACACACGGTT 1576
101 euGlyValTyrSerProHisPro..... 108
1577 CGGCTAATTTGGCTGCCATCCCACTGAAAGCCGAGCCGACGACACC 1626
108 ..... 108
1627 CACACCTCTCTCTGCAAGCCGCGAGCAAAAAGTACTTTGGATGCCAA 1676
109 .....TyrGluGlnGluValS 114
1677 CCTGTATCAAGAACAATTACGAAATTTCTGTCTGAAGACACAAAGCTGA 1726
114 eAlaLeuGlyGlyLe..... 119
1727 CCACACGCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
120 .....ProTyrSerGlnIleTyrGlyTyrTyrArgAlaHisnheG 133
1777 TGTAGCAATCTTATTCGAAAGAACTGGAATGGCAGATGCAAAAT..... 1821
133 yValIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrT 150
1822 ....ATGACAAAGCCCAAGATCCCGGTATGACAGCTGACGCGGCGCT... 1863
150 yraspsplGluAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 166
1864 .....CTGAATGTGAC.....AAAGTACGCTGT 1887
167 PheProProAspHisGlnAlaTPrArgGluGluProTrpIleHisHsAl 183
1888 TTGTGTCT.....GAGCGCTGGAAACTGTTCGG 1916
183 aProGlnGlyCysGlyAsnSerArgTyrIleThrGlyAspThrCysA 200
1917 CTGCTGGGTATGCGAAAGCAAA.....CTGCGGCGACAAACAGCC 1960
200 snGluGlnThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 216
1961 TGCTGTCCACCCAGCCGTGAAGATGTCGCCGTATCGACTATGAAGT 2010
216 ..... 216
2011 CCGAGCGAAAAATGGGGCTGTCTCCCGCTGACTATCTGGCGCGAA 2060
217 .LysValIysArgGlnIlePheSerAspTyrGlnSerGlu.....V 230
2061 AAAGGTCAAAAGACGCGCAATACACCGTTTATGAACAAACAGCGCTGGG 2110
230 alAspIleTyrAsnArgIleArgAsp 238
:: .....

```


1727 CTAAAGGCTAAATCTTTTAAAT
|||::|||::|||:::
1747

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-408-095-26

seq_documentation_block:

Sequence 6, Application US/08/408095
Patent No. 5858678
GENERAL INFORMATION:
APPLICANT: Chinadurai, Govindaswamy
TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,095
FILING DATE: 21-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO.: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 26..505
IS-08-408-095-26

```
alignment_scores:
  quality: 75.00
  ratio: 1.19
  percent_similarity: 41.615
  length: 161
  gaps: 6
  percent_identity: 21.118
```

```

alignment_block:
US-09-528-682-1 x US-08-408-095-26/rev ..
Align seg 1/1 to reverse of: US-08-408-095-26 from: 1 to: 946

```

```

100 valLeuGIyValIySerProHis..... 107
    ::::::::::::::::::::
882 CTCCTCGAATCTATTCAGATAGATCTCTGGCAGGCGCATCATATCA 833
108 .....ProTyrGIuGIuValSerAlaLeuG 117
    ||| ||| |||
832 CACAGGGGTGCACAGCAGCATCCCCCTGCATACCTGCCGTGTGGCAGGG 783
117 IyGIyIleProTyrSerGIuIleTyrGIyTyrTyrTyrValIaIaPheGIy 133
    ::::: ||| ::::: ::::: :::::
782 CAGAGGTCAATGGCTCT.....TCTGGGCACAGCCACAGTTAAC 742
134 ValIleAspGIuATyGLeuHisATyAsnArgGIuTyrArg..... 146
    ::::: ||||| ||| ::::: |||
741 CGGTGCACATTTGCAGAGCCATGAGAAATTCATCTTTTGAAAAAACGAAA 692
147 .....AspATyTyrTyrATyAsnLeuAsnIleAlaProAlaGIuAspG 161
    ::::: ::::: ||||| ||| |||
691 CTGCATTAATAAAAAACAAAAAACCCTGATATTAACCCTCAGCAGTGT 642
161 IyTyrArgLeuAlaGIyPhe..... 167
    ::::: |||||
641 CCAGCATATCTCGGGGTTTAATATAAAAAAGCATCATGAGAAACAGT 592
168 .....ProProAspHisGIuAlaTyr.. 174
    ||||| ::::: |||||
591 TAAAAAGATTAACAGCAGCAGGCCCCACCTCCAGCGGATGGTCATGGGG 542
175 .....ArgGIuGIuProTyrIleHisIleAlaProGIuIyC 187
    ||| ||||| ||| :::::
541 GTGGGGCCAGCCAGCCCTGAGCCGCCGGGGCCCTCACTTGAAGCAGAGT 422
187 ys.....GlyAsnSerSerATyGHIle 194
    ||| ::::: ||||| :::::
491 GCAGCCCCCGCTGAGCAGCGGACAGACAGCCACAGCAGCAGCAGCAGC 442
195 ThrGIyAspTyrCysAsnGIuGIuThrGIuAsn 205
    ::::: ||||| ::::: ||||| :::::
441 GCCACGACACACTGTGTTCAGAGACACCCAGCAGC 409

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-182-117-1

seq_documentation_block:

Sequence 1, Application US/09182117
 Patent No. 6204436
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Transgenic Plant
 NUMBER OF SEQUENCES: 27
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/182,117
 FILING DATE:
 CLASSIFICATION:
 INFORMATION FOR SEQ ID NO: 1:
 .SEQUENCE CHARACTERISTICS:
 LENGTH: 8012 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 OS-09-182-117-1

alignment_scores:

Quality:	74.50	Length:	102
Ratio:	1.263	Gaps:	4
Percent Similarity:	57.843	Percent Identity:	24.510

alignment_block:

US-09-528-682-1 x US-09-182-117-1/rev ..

Align seg 1/1 to reverse of: US-09-182-117-1 from: 1 to: 8012

```
74 GluSerIleLeuSerGIYrSerThrTYrTlleTYrValIleAlaTh 90
||| ::||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
5175 CAGTTTCTCCTGCCAGGCCAGAACTTCTTTCCAGTACCTTCTCTGCCG 5126
90 rAlaPro...AsnMetPheAsnValAsnAspValIleuGIYValTYrSerP 106
5125 TTTCGAATGCGCGCTTTGGACATACCATCCCGTAAATACGGTTCCAGGCAC 5076
106 rOHISProTYrGIuGIuValSerAlaLeuGIYIleProTYrSer 122
5075 AGCAC.....ATCAAGAGATCGCTGATGCTATCGGTGAGCGTCGCA 5032
123 GluIleTYrGIYrPYrArGValAsnPheGIYValIleAspGIuArGLe 139
5031 GAACATTAC.....ATTGACGACAGTGATCGAGCGCTCGGGT 4994
139 uHISArgAsnArgGIuTYrArGAspArGYrTYrArGAsnLeuAsnIleA 156
4993 CGAGTTTACCGGTGCTTCGCCAG.....TGCGAAATGCCGAATAT 4950
156 IaProAlaGIuAspGIYrTYrArGLEuAlaGIYrPheProAlaSPHISGI 172
4949 TCCGTCACAGACGGGTATCCGGTTCGTGGCAATACTCCACATCACCCAC 4900
173 AlaTrp 174
|||||
4899 GCTTGG 4894
```


[illegible]

/cgn2_6/p.todata/2/pna/US0951.COMB	seq:US-09-534-859-383	-	80.00	94.94	7.7e+04	848	/cgn2_6/p.todata/2/pna/PCRTUS.COMB	seq:PCRT-US01-08631-14949	+	78.50	124.40	1.9e+00	328
/cgn2_6/p.todata/2/pna/US0951.COMB	seq:US-09-803-738-383	-	80.00	94.94	7.7e+04	848	/cgn2_6/p.todata/2/pna/US099A.COMB	seq:US-09-902-540-6744	+	78.50	123.52	2.0e+03	328
/cgn2_6/p.todata/2/pna/US0951.COMB	seq:US-09-534-859-382	-	80.00	93.76	9.0e+04	952	/cgn2_6/p.todata/2/pna/US099A.COMB	seq:US-09-902-540-554	+	78.50	121.83	2.5e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-803-738-382	-	80.00	93.76	9.0e+04	952	/cgn2_6/p.todata/2/pna/US096A.COMB	seq:US-09-60-068-175-465	-	78.50	121.70	2.5e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-328-352-493	+	79.50	149.40	71.67	359	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-08-961-527-142	-	78.50	120.74	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	144.77	129.81	58	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-654-617-21013	-	79.50	143.69	149.13	6	/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-60-029-960-45	-	78.50	120.71	2.8e+03	328
/cgn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-6												

[illegible][illegible]

[illegible][illegible]

[illegible]

/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-111 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-112 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-113 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-114 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-115 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-116 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-117 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-118 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-119 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-120 + 75.00 105.75 1.9e+04 108
/cgn2_6/ptodata/2/pna/US6020.COMB.seq:US-60-205-419-121 + 75.00 105.75 1.9e+04 108
seq_name: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:US-09-724-315-4
seq_documentation_block:
; Sequence 4, Application US/09724315
; GENERAL INFORMATION:
; APPLICANT: HAINES, Joel R.
; APPLICANT: ARLINGTON, Joshua
; TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
; FILE REFERENCE: APE41
; CURRENT APPLICATION NUMBER: US/09/724,315
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5500
; TYPE: DNA
; ORGANISM: plasmid pUV2004
US-09-724-315-4

alignment_scores:
Quality: 1297.00 Length: 240
Ratio: 5.404 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.167
alignment_block:
US-09-528-682-1 x US-09-724-315-4 ..
Align seg 1/1 to: US-09-724-315-4 from: 1 to: 5500

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3975 AATGGCGACAATATATACCTGCTGACTCTAGACCCCAATGAAATATA 4024
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4025 ACGTCCGAGGCTTATATGCGCCAGAGGCGAATATGAGTACTTCATAGAG 4074
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4075 GAACCTAAATGATATTAATCTTATATGATCGCGAGAGGAGCAACAAC 4124
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4125 GGCCTTGTCAATATGATGACGAGATATGTTCCACTTCTCTAGTTGAG 4174
67 gSerIleHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4175 AAGTCTCACTTAGAGGACAGCTATATATTCAGGATATTCACCTTACT 4224
84 yTlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4225 ATATATATATATAGCGACGACCAATATGTTAATGTTAATGATGTA 4274
101 LeuGlyValIleTyrSerProHisProTyrGlnGlnIleValSerAlaLeu 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4275 TTAGCGGATAGACGCGCTCACCAATATGACAGAGGTTCTGCGTTAG 4324
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4325 TGGATACCAATATCTCAGATATATGATGATGCTATCGTATTTTGGTG 4374
```

```
134 lIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4375 TGATGATGTAACCATTTACATCGTATACAGGAAATATAGACCGGTTTAC 4424
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4425 AGAAATCTGAATATAGCTCCGCGACAGAGATGCTTACAGATTACAGTTT 4474
167 eProProAspHisGlnAlaIlePheArgGluIleProTyrPheHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4475 CCAACCGATCCACCAAGCTTGAGAGAGAACCCCTGATTCATCATGACAC 4524
184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4525 CACAGGTTCTGGAATTCATCATCAAGAACATTTACAGTGTACTTGTAT 4574
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4575 GAGGAGACCCAGAACTCGACACAAATATATCTCAGGAAATATCAATCAA 4624
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4625 AGTTAAGAGGACAGATATTTTCAGACTATCTCAGAGGTTGACATATATA 4674
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4675 ACAGAAATTCGGATGAAATTA 4694
```

seq_name: /cgn2_6/ptodata/2/pna/US099B.COMB.seq:US-09-950-335A-5
seq_documentation_block:
; Sequence 5, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 5
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-950-335A-5

alignment_scores:
Quality: 1293.00 Length: 240
Ratio: 5.410 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 99.167
alignment_block:
US-09-528-682-1 x US-09-950-335A-5 ..
Align seg 1/1 to: US-09-950-335A-5 from: 1 to: 723

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 AATGGCGACAATATATACCTGCTGACTCTAGACCCCAATGAAATATA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 ACGTCCGAGGCTTATATGCGCCAGAGGCGAATATGAGTACTTCGATAGAG 100
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 GAACCTAAATGATATTAATCTTATATGATCACCAGAGAGAACACAAC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```



```

151 GGCTTTGTCAGATATGATGACGATATGTTCCACTTCTTCTTGTAGTTGAG 200
67 gserAlaHsLeuAlaGlyInSerIleuSerGlyTyrrSerThrTyT 84
|||||
201 AGTGTCTCACTTAGCAGACAGCTATATATATCAGGATATTCACACTTACT 250
84 yrlleTyValIleAlaThrAlaProAsnMePheAsnValAsnAspVal 100
|||||
251 ATATATATGTATTAGGACAGACCAATATGTATTATGTATATGATGTA 300
101 LeuGlyValTyrrSerProHsProTyrrGluGluGluValSerAlaLeuG 117
|||||
301 TTAGGGGTATACAGCCCTCACCATTGTGAACAGAAAGTTTCCGCTTAGG 350
117 yGlyIleProTyrrSerGlnIleTyrrGlyTrpTyrrArgValAsnPhGlyV 134
|||||
351 TGGAAATCAATATCTCAGATATATGATGATGTAATCGTTAAATTTGGTG 400
134 AlIleAspGluArgLeuHsIleArgAsnArgGluTyrrArgAspArgTyrr 150
|||||
401 TGATTGATGACGATTACATGGTAACAGGAAATATAGACACCGGTATTAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPh 167
|||||
451 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACAGATTAGCAGGTTT 500
167 eProProAspHsIleAlaTrpArgGluGluProTrpIleHsIleAlaP 184
|||||
501 CCACCCGGATCCACCAAGCTTGGAGAGAACCCCTGATTCATCATCAGCAC 550
184 roGlnGlyScGlyAsnSerSerArgTrIleThrGlyAspThrCysAsn 200
|||||
551 CACAAGGTTGGAAATTCATCAGAAACAATCAGACGTATCTTAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrrLeuArgGluTyrrGlnSer 217
|||||
601 GAGGAGACCCAGAAATCTGAGCACAAATATCTCAGGAAATATCAATCAAA 650
217 sValIysArgGlnIlePheSerAspTyrrGlnSerGluValAspIleTyrr 234
|||||
651 AGTTAGAGGACAGATATTTTCAGACTATCATCAGTCAGAGTTGACATATATA 700
234 snArgIleArgAspGluLeu 240
|||||
701 ACAGAAATTCGGGATGAAATTA 720

```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-3

seq_documentation_block:

Sequence 3, Application PC/TUS9930747

GENERAL INFORMATION:

APPLICANT: Boyce Thompson Institute for Plant Research at Cor

TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in

FILE REFERENCE: 4868/85427

CURRENT APPLICATION NUMBER: PCT/US99/30747

CURRENT FILING DATE: 1999-12-22

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn Ver. 2.1

Seq ID NO 3

LENGTH: 782

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (3)..(782)

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:E. coli

OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize

OTHER INFORMATION: expression in plants.

PCT-US99-30747-3

alignment_scores:

Quality:	1287.00	Length:	240
Ratio:	5.362	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	98.333

alignment_block:

US-09-528-682-1 x PCT-US99-30747-3 ..

Align seg 1/1 to: PCT-US99-30747-3 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrrArgAlaAspSerArgProProAspGluIle 17
|||||
60 AACGGAGACAAGCTCTATAGGGCTGACTAGGCCCCCGCAGATAGATCA 109
17 sArgSerGlyGlyLeuMetProArgGlyIleAsnGlyTyrrPhAspArg 34
|||||
110 GAGGTGTGGAGGTCTCATGCCAAGGGGAGACAAATGAGTACTTTGATA 159
34 IYThrGlnMetAsnIleAsnLeuTyrrAspHsIleAlaArgGlyThrGln 50
|||||
160 GAACCCAAATGAACATCAACCTTATGACCATCTAGGGGAACCCAACT 209
51 GlyPheValArgTyrrAspAspGlyTyrrValSerThrSerLeuSerLeu 67
|||||
210 GGCTTTGTGAGTATGATGATGATGATGATGTTCCACCTTCCCTTAGT 259
67 gserAlaHsLeuAlaGlyInSerIleuSerGlyTyrrSerThrTyT 84
|||||
260 GTCTGTCTCACTTGGCAGCAACATCCCTCAGATATCTCCACCTACT 309
84 yrlleTyValIleAlaThrAlaProAsnMePheAsnValAsnAspVal 100
|||||
310 ACATATATGTGATGTCTACAGCACCAACATGTTCATATGTAATGATG 359
101 LeuGlyValTyrrSerProHsProTyrrGluGluGluValSerAlaLeu 117
|||||
360 TTGGGAGTGATAGCCCTCACCATATGAGCAAGAGGTGTCTGCTTGG 409
117 yGlyIleProTyrrSerGlnIleTyrrGlyTrpTyrrArgValAsnPhGlyV 134
|||||
410 TGGAAATCCCATCTCCCAATCTATGATGATGTAAGGTTGAACCTTGG 459
134 AlIleAspGluArgLeuHsIleArgAsnArgGluTyrrArgAspArgTyrr 150
|||||
460 TGATTGATGAGAGGCTCCATAGGAATAGGAGATATAGGACAGGTACT 509
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPh 167
|||||
510 AGGAACCTCAACATAGCTCCAGCAGAGAGATGTTATAGTTGGCAGGTT 559
167 eProProAspHsIleAlaTrpArgGluGluProTrpIleHsIleAlaP 184
|||||
560 CCACACGACACCAACAGCTTGGAGGAGAGGCCCTGATTCACACATGCA 609
184 roGlnGlyScGlyAsnSerSerArgTrIleThrGlyAspThrCysAsn 200
|||||
610 CACAGAGTTGTGAGACTCTCTCAAGCAACATCACAGGTGACACTTCA 659
201 GluGluThrGlnAsnLeuSerThrIleTyrrLeuArgGluTyrrGlnSer 217
|||||
660 GAGGAGACCCCAAACTTATAGCACCATCTACCTTAGAAGTAAACCAAT 709
217 sValIysArgGlnIlePheSerAspTyrrGlnSerGluValAspIleTyrr 234
|||||
710 GTGAGAGAGCAATCTTCTCAGACTACCAATCAGAGTGGACATCTACA 759
234 snArgIleArgAspGluLeu 240
|||||
760 ATAGGATTAGGAATGAACTC 779

```

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-3

seq_documentation_block:


```

; Sequence 3, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arntzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-3

```

```

alignment_scores:
    Quality: 1287.00      Length: 240
    Ratio: 5.362          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 98.333

```

alignment_block:

US-09-528-682-1 x US-09-470-124-3 ..

Align seg 1/1 to: US-09-470-124-3 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
  |||||
60 AACGGAGACAAAGCTCTATAGGCGCTAGCGCCCGCCAGATGAGATCAAA 109
  |||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  |||||
110 GAGGTCTGGAGGTCTCATGCCAAGGAGCACAAATGAGTACTTGTATAGG 159
  |||||
34 LThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||||
160 GAACCCAAATGAAATCATCACTCTATGACCATGCTAGGGAGACCCAACT 209
  |||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||||
210 GGCCTTGTGAGGTATGATGATGATGTGTCCACCTCTAGCTTGTGAG 259
  |||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||||
260 GTCTGCTCAGCTTGGCAGGACATCCATCTCTCGGATACCTCCACTACT 309
  |||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||||
310 ACATCTATGATGATGTGTACAGACACAAACATGTTCAATGTGAATGAGCTG 359
  |||||
101 LeuGlyValIleTyrSerProHisProTyrGluGlnIleValSerAlaLeu 117
  |||||
360 TTGGAGATGATATACCCCTCACCCATATGAGCAAGAGGTGTCTGCTTTGG 409
  |||||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
  |||||
410 TGGATCCCACTCCCAAAATCTATGATGATGATGATGATGATGATGATG 459
  |||||
134 aIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||||
460 TGATTGATGAGAGGCTCCATATGAAATAGGGGATATAGGACAGGTACTAT 509
  |||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
  |||||

```

```

510 AGGAACCTCAACATAGCTCCAGCAGAGATGTTATAGCTTGGACGTTT 559
  |||||
167 eProProAspHisGlnAlaThrArgGlnGluProTyrPheHisAlaP 184
  |||||
560 CCCACCAAGACCAAGCTGGAGGAGAGCCCTGATGCCACCATGAC 609
  |||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||||
610 CAAAGGTGTGAGAGCTCCTCAAGCAGCATCAGAGGTGACACTTGCAAT 659
  |||||
201 GluGlnThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
  |||||
217 sValTyrArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||||
710 GGTGAAGAGCAATCTTCTCAGACTACCAATCAATGAGGTGACATCTACA 759
  |||||
234 snArgIleArgAspGluLeu 240
  |||||
760 ATAGATTTAGATGATGAATC 779
  |||||

```

seq.name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-315-6

seq_documentation_block:

```

; Sequence 6, Application US/09724315
; GENERAL INFORMATION:
; APPLICANT: HAYNES, Joel R.
; APPLICANT: ARRLINGTON, Joshua
; TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
; FILE REFERENCE: AFP41
; CURRENT APPLICATION NUMBER: US/09/724,315
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5488
; TYPE: DNA
; ORGANISM: plasmid pPUV2007
US-09-724-315-6

```

```

alignment_scores:
    Quality: 1269.00      Length: 236
    Ratio: 5.400          Gaps: 0
    Percent Similarity: 99.576    Percent Identity: 98.729

```

alignment_block:

US-09-528-682-1 x US-09-724-315-6 ..

Align seg 1/1 to: US-09-724-315-6 from: 1 to: 5488

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
  |||||
3975 AATGGGACAAATTAATTAATCTTATGATGATGATGATGATGATGATG 4024
  |||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  |||||
4025 ACCTTCCGAGAGCTTATGCGCATGCGCATATGATGATGATGATGATG 4074
  |||||
34 LThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||||
4075 GAACCTAAATGATATTAATCTTATGATGATGATGATGATGATGATG 4124
  |||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||||
4125 GGCCTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATG 4174
  |||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||||
4175 AAGTCTCAGCTTACGAGACAGCTCTATATATATATATATATATATAT 4224
  |||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||||

```



```

|||||
4225 ATATATATGTTATGACGACACCAATATGTTATGTTAATGATGTA 4274
101 LeuGlyValIYrSerProHISProTYrGluGlnIuValSerAlaLeuG 117
4275 TTGGCCCTAATACGCCCTCACCATAATGAACAGAGGTTCTCGGTAGG 4324
117 YGlyIleProTYrSerGlnIleTYrGlyTYrPYrArgValAsnPhgIy 134
4325 TGGAAATCCATATTTCTCAGATATATGATGATCGTGTATTAATTTGGTG 4374
134 allleAspGluArgLeuHISArgSnaArgLUrTYrArgAspArgTYrTYr 150
4375 TGATGTATGAACGATTCATCGTAAACAGGGAATATAGACCGGTATTTAC 4424
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPh 167
4425 AGAAATCTGAATATATAGCTCCGCCAGAGATGTTACAGATTACAGGTTT 4474
167 eProProAspHISGlnAlaTrpArgGluGluProTrpIleHISHisAlaP 184
4475 CCCACCGGATCACCAGCTTGAGAGAGAAACCTGGATTCATCATGACAC 4524
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
4525 CACAAGGTGTGGAAATTCATCAAGAACAAATTACAGGTGATCTGTAAT 4574
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerLy 217
4575 GAGAGAGCCAGATCTGACACAAATATATCTCAGGAAATATCAATCAA 4624
217 SValIysArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
4625 AGTTAAAGCAGCATTTTTCAGACTATCAGTCAGAGGTTGACATATATA 4674
234 SnaArgIle 236
4675 ACAGAAATT 4682

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-5
seq_documentation_block:
: Sequence 5, Application us/08256003
: GENERAL INFORMATION:
: APPLICANT: Domenighini, Mario
: APPLICANT: Rappuoli, Rino
: TITLE OF INVENTION: Immunogenic Detoxified Mutants of
: TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt. Their Preparation and
: TITLE OF INVENTION: Their Use for the Preparation of Vaccines
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/256,003
: FILING DATE: 11-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0315.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708

```

```

: TELEFAX: (510) (655-3542)
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: NAME/KEY: CDS
: LOCATION: 1..711
: US-08-256-003-5

alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:
US-09-528-682-1 x US-08-256-003-5 ..

Align seq 1/1 to: US-08-256-003-5 from: 1 to: 711

1 AsnGlyAspArgLeuTYrArgAlaSpSerArgProProAspGluIleLY 17
|||||
1 AATGGCGACGATTTATACCTGCTGACTTACGACCCCGAGATGAATAA 50
17 sArgSerGlyLeuMetProArgGlyHISAsnGluTYrPheAspArg 34
|||||
51 ACFTTCCGGAGCTTTATGCCAGAGGT...AATGAGTACTTCGATGAG 97
34 LYThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
98 GAACCTCAATGAATATTAATCTTATGATCAGCGAGAGGAACACAAAC 147
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuArg 67
|||||
148 GCGTTTGTACGATATGATGACGCGATATGTTCCACTTCTCTAGTTTGGAC 197
67 gSerAlaHISLeuAlaGlnIleSerIleLeuSerGlyTYrSerThrTYrT 84
|||||
198 AAGTGCACACTTACGACGACGATATATATATATCAAGATATCACTTACTA 247
84 YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
|||||
248 TATATATCTGTTATAGCA.....AATATGTTAATGTTAAATGATGTA 288
101 LeuGlyValIYrSerProHISProTYrGluGlnIuValSerAlaLeuG 117
289 ATTAGCGTATACAGCCCTCACCCATATGACAGAGGTTTTCGCTTAGG 338
117 YGlyIleProTYrSerGlnIleTYrGlyTYrPYrArgValAsnPhgIy 134
|||||
339 TGGAAATCCATATTTCTCAGATATATGATGATCGTGTATTAATTTGGTG 388
134 allleAspGluArgLeuHISArgSnaArgLUrTYrArgAspArgTYrTYr 150
|||||
389 TGATGTATGAACGATTCATCGTAAACAGGGAATATAGACCGGTATTTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPh 167
|||||
439 AGAAATCTGAATATATAGCTCCGCCAGAGATGTTACAGATTACAGAGTTT 488
167 eProProAspHISGlnAlaTrpArgGluGluProTrpIleHISHisAlaP 184
|||||
489 CCCACCGGATCACCAGCTTGAGAGAGAAACCTGGATTCATCATGACAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
539 CACAAGGTGTGGAAATTCATCAAGAACAAATCAACAGGTTGATCTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerLy 217

```



```

|||||
589 GAGGAGACCCAGAACTGACACATATATCTCAGGGAATATCAATCAA 638
217 svaIIysArgInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
   |||||||
639 AGTTAAGAGGAGATATTTTCAGACTATCAGTCAGAGGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
   |||||||
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-044-696-1

```

seq_documentation_block:
: Sequence 1, Application US/09044696
: GENERAL INFORMATION:
: APPLICANT: BARCHELD, GAIL.
: APPLICANT: DEL GIUDICE, GIUSEPPE
: APPLICANT: RAPPUOLI, RINO
: TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
: TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
: ADDRESS: R440
: STREET: P.O. BOX 8097
: CITY: EMERYVILLE
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94662-8097
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/044,696
: FILING DATE: 18-MAR-1998
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,227
: FILING DATE: 21-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: HARBIN, ALISA A.
: REGISTRATION NUMBER: 33,895
: REFERENCE/DOCKET NUMBER: 1393.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 655-8730
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..708
: US-09-044-696-1

```

alignment_scores:

```

Quality: 1183.00      Length: 240
Ratio: 5.189          Gaps: 2
Percent Similarity: 95.000      Percent Identity: 93.333

```

alignment_block:

US-09-528-682-1 x US-09-044-696-1 ..

Align seg 1/1 to: US-09-044-696-1 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLys 17

```

|||||
1 AATGCCGACAGATATATACCGTGCTGACTCTAGACCCACAGATGAATAAAA 50
17 sArgSerGlyGlyLeuMetProAlaGlyHisAsnGluTyrPheAspArg 34
   |||
51 ACGTTTCCGAGCTTATGCCACAGAGT...AAGAGTACTTTCGATAGAG 97
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
98 GAATCAATGAATATTTATCTTATGATCAGCGGAGAGACAAAC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
148 GCGTTTGTACAGATATGATGACGCGATATGTTCCACTCTCTAGATTGAG 197
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyr 84
198 AAGTCTCACTTAGCAGACACGATATATATATATATATATATCTACTTA 247
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATGCA.....AATATGTTATATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
289 ATTAGCGTATACAGCCCTCACCCATATGAAACAGAGGCTTCTGCGTTAG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
339 TGGAAATACCATTTCTCAGATATATGATGATGATGATGATGATGATG 388
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
389 TGATGATGTAACGATTTACATCGTACAGGGAATATACAGCCGATATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
439 AGAAATCTGAATATAGCTCCGCGCAGAGATGCTTACGATTAGAGGTTT 488
167 eProProAspHisGlnIleAlaThrArgGluGluProTyrIleHisIleAla 184
489 CCCACCGGATCCACAGCTTGAGAGAGAAACCTGATTCATCATCATC 538
184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAGAGTTGTGGAGATTCAACAGAAACATACAGGTGATCTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
589 GAGGAGACCCAGAAATCTGACACACATATATCTCAGGGAATATCAATCAA 638
217 svaIIysArgInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
639 AGTTAAGAGGAGATATTTTCAGACTATCAGTCAGAGGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-7

seq_documentation_block:

```

: Sequence 7, Application US/08256003
: GENERAL INFORMATION:
: APPLICANT: Domenighini, Mario
: APPLICANT: RAPPUOLI, RINO
: APPLICANT: PIZZA, Mariagrazia
: TITLE OF INVENTION: Immunogenic Detoxified Mutants of
: TITLE OF INVENTION: Cholera Toxin and of the Toxin B, Their Preparation and
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLys 17

STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256.003
FILING DATE: 11-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..723
US-08-256-003-7

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x US-08-256-003-7 ..

Align seg 1/1 to: US-08-256-003-7 from: 1 to: 723

1 AsgGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
1 AAAGATGATAGTATATCGCGCATCTAGACCTCCATGATAAAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 GCAGTCAGGCGGTCTTATGCCAAGAGACAGAGTAGTACTTTGACCGAG 100
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTAGCTCAAAAGATATCAACCTTATGATCATGCAAGAGAGACTCAAGAC 150
51 GlyPheAlaArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
151 GGATTTGTTAGSCAGCATGATGATGTTTCCACCTCAATTAGTTTGGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTGGCCCATTTAGTGGTCAACATATATGTCTGTCATCTTACTTATTT 250
84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCACATGCACCCAAACATGTTTAACTTAATGATGA 300
101 LeuGlyValAlaTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
301 TTAGGGCATACAGTCTCTCATCCAGATGAACAAGATTTCTGCTTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhcGly 134

|||||
351 TGGGATTCCATFACTCCCAAAATATATGATGATATCGAGTTTCATTTGGGG 400
134 allleaspgluatrgleuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
401 TGCTTGATGAACAATATCATCTGTAATAGGCGCTACAGATAGATATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
451 AGTACTTATGATATGTCTCCAGCAGCAGATGTTATGATGATGGCAGGTTT 500
167 eProProAspHisGlnAlaIleTyrArgGluGluProTyrPheHisAlaP 184
501 CCTCCGCGACATAGAGCTTGAGGAGAGACCGGTGATTCATCATATCAC 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CCGCGGCTGTGGGAATGCTCAAGATCATCATCATATCTACTTCCGAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuAlaArgGluTyrGlnSer 217
601 GAAAGAACCCAAAGTCTAGGTGTAATAATCTTGACGAATACCAATCTAA 650
217 sVallysArgGlnIlePheSerAspTyrGlnSerGluValAlaAspIleTyr 234
651 AGTTAAAGACAAATATTTTCAGGCTATCATCTGATATGATATACACATA 700
234 snArgIleArgAspGluLeu 240
701 ATAGAAATTAAGATGAATTA 720

seq_name: /cgn2_6/plodata/2/pna/US090_COMB.seq:US-09-044-696-3

seq_documentation_block:

Sequence 3, Application: US/09044696

GENERAL INFORMATION:

APPLICANT: BARCHELD, GAIL

APPLICANT: DEL GIUDICE, GIUSEPPE

APPLICANT: RAPPOLO, RINO

TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL

NUMBER OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -

ADDRESSEE: R440

STREET: P.O. BOX 8097

CITY: EMERYVILLE

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/044,696

FILING DATE: 18-MAR-1998

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/041,227

FILING DATE: 21-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: HARBIN, ALISA A.

REGISTRATION NUMBER: 33,895

REFERENCE/DOCKET NUMBER: 1393.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 655-8730

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 723 base pairs

TYPE: nucleic acid


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..720
; NAME/KEY: CDS
; LOCATION: 1..720
; US-09-044-696-3

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent Similarity: 94.167  Percent Identity: 81.667

alignment_block:
  US-09-528-682-1 x US-09-044-696-3  ..

Align seg 1/1 to: US-09-044-696-3 from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleTyr 17
1 AATGATGTAAGTATATATCGGACGATTCCTGATCGAATGAAATGAA 50
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 GCAGTCAGGTGGTCTTATGCCAAGAGACAGATGATCTTGACCGAG 100
34 LThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACCAATGAAATATCAACCTTATGATCATGCAAGAGAACACAGACG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
151 GGATTTGTTAGGCACGATGATGATGATGTTCCACCTCAATGATTTGAG 200
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTCCCACTTAGTGGGTCAACATATATGTCGTGCTACTTACTTAT 250
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCACTGACCCCAACATGTTAACTTAATGATGTA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCATCCAGATGACACAGAGTTTCTGCTTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
351 TGGGATTCATACCTCCCAATATATGATGATGATGATGATGATGATG 400
134 alIleAspGluArgLeuHisArgAsnAlaGluTyrArgAspArgTyrTyr 150
401 TGCCTGATGAACAATATCATGCTATATAGGGCTCAGAGATGATATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
451 AGTAACTTAGATATTGCTCCAGCAGCAGATGTTATGATGATGATGAT 500
167 eProProAspHisGlnAlaTyrArgGluGluProTyrPheIleHisAlaP 184
501 CCCTCCGAGCATAGAGCTTGAGAGAGAGCCGCTGATCATCATGAC 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CGCGGGGTTGGGAAATGCTCCAGATCATGATCAGTAATATCTGCGAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
601 GAAATAACCCAAAGCTAGGTAAATTCCTTGACGATACCAATCTTAA 650
```

```
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAlaSerIleTyr 234
651 AGTAAAGACAAATATTTCCAGGCTATCATCTGATGATGATGATGATGAT 700
234 snArgIleArgAspGluLeu 240
701 ATAGAAATTAAGATGAAATTA 720

seq_name: /cgn2_6/plodata/2/pna/US099B_COMB.seq:US-09-950-335A-1

seq_documentation_block:
; Sequence 1, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Vibrio cholerae
; US-09-950-335A-1

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent Similarity: 94.167  Percent Identity: 81.667

alignment_block:
  US-09-528-682-1 x US-09-950-335A-1  ..

Align seg 1/1 to: US-09-950-335A-1 from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleTyr 17
1 AATGATGTAAGTATATATCGGACGATTCCTGATCGAATGAAATGAA 50
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 GCAGTCAGGTGGTCTTATGCCAAGAGACAGATGATCTTGACCGAG 100
34 LThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACCAATGAAATATCAACCTTATGATCATGCAAGAGAACACTCAGACG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
151 GGATTTGTTAGGCACGATGATGATGATGTTCCACCTCAATGATTTGAG 200
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTCCCACTTAGTGGGTCAACATATATGTCGTGCTACTTACTTAT 250
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATAGCCACTGACCCCAACATGTTAACTTAATGATGTA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCATCCAGATGACACAGAGTTTCTGCTTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
351 TGGGATTCATACCTCCCAATATATGATGATGATGATGATGATGATGAT 400
134 alIleAspGluArgLeuHisArgAsnAlaGluTyrArgAspArgTyrTyr 150
```



```
401 TCGTGTGATGACATTACATGTAATAGGGGCTACAGAGATAGATTATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuIleAlaGly 167
451 AGTAAGTATGATATGCTCCAGCAGCAGATGATGATGATGATGATGATGATG 500
167 eProProAspHisGlnAlaIleAlaTyrArgGluGluProTyrIleHisIleAla 184
501 CCCTCCGAGACATAGACCTTGAGAGGAGAGCCGTGATTCATCATCATCAG 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CCGCGGCTTGCGGAATGCTCAAGATCATCATCATCATCATCATCATCATCAT 600
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIly 217
601 GAAAAAACCAGAGCTGAGGTGTAATAATCTTGACGAAATACCAATCTAA 650
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
651 AGTTAAAGCAAAATATTTTCAGGCTATCATCTGATATGATGATGATGATG 700
234 snArgIleArgAspGluLeu 240
701 ATAGAATTAGATGATGATTA 720

seq_name: /cgn2_6/ptodata/2/pna/PCIVUS_COMB.seq: PCT-US99-30747-45

seq_documentation_block:
; Sequence 45, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 45
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-45

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x PCT-US99-30747-45 ..

Align seq 1/1 to: PCT-US99-30747-45 from: 1 to: 777

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIly 17
111 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 104
55 AATGATATGACAGCTCTATAGGGCAGACTCAAGACTCTCTGATGATGATCAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
105 GCATTCAGGTGGTCTTATGCCAGAGGACATCTGAGTACTTGTACAGGG 154
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
155 GTACTAGATGAAACATCAACCTTTATGACCATGCAAGGGAACATCAACT 204
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeuArg 67
205 GGATTTGTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATG 254
67 gSerAlaHisIleuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
255 GTCTGCCCATCTGGTGGTCAAACTATCTCTCTGCTGACCTTACTTACT 304
84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
305 ACATCTATGTGATGATGCCATGACCCCAACATGTTCAATGGAATGATGATG 354
101 LeuGlyValIlySerProHisProTyrGlnGlnGluValSerAlaLeuG 117
355 TTGGAGACATACACCCCTCACCAGATGACCAAGAGGTGCTGCTGGTGGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
405 TGGAAATCCCATCTCCCAATCTATGATGATGATGATGATGATGATGATG 454
134 alIleAspGluArgLeuGlnHisArgAsnArgGluTyrArgAspArgTyr 150
455 TCGTTGATGAGCAACTCCATAGAAATAGGGGCTACAGGATGATGATGAT 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
505 AGCAACTTGGACATTTGCTCCAGCAGCAGAGTGTATGATGATGATGATG 554
167 eProProAspHisGlnAlaIleAlaTyrArgGluGluProTyrIleHisIleAla 184
555 CCCCTCCAGAGCATAGGGCTTGAGAGGAGAGCCCTTGATTCACCATGAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
605 CACCAGGTGTGGAATGCTCCAGAGGTCAAGATGACAGACACACACTTGAT 654
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIly 217
655 GAAAAAGACCCAAATCTTGGGTGTAAGTCTCTGATGATGATGATGATGAT 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
705 GGTAAGAGGCAAAATCTTTCAGGCTATCATCTGATGATGATGATGATG 754
234 snArgIleArgAspGluLeu 240
755 ATAGATCAAGATGATGATC 774

seq_name: /cgn2_6/ptodata/2/pna/US07_COMB.seq: US-07-694-733-1

seq_documentation_block:
; Sequence 1, Application US/07694733
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: Recombinant DNA-derived
; TITLE OF INVENTION: Cholera Toxin
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1840 Dehavenland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh OS 7.0.
; SOFTWARE: Microsoft Word Version 5.0
```



```

655 GAAAAAACCCAAAGCTAGTGTAATAATTCCTTGACGAATACCAATCTAA 704
217 STAlYsArGinGInIePheSeRAsFTyGInSeRGIuValAsPlIeTyra 234
|||||.....|||
705 AATYTAAGACCAAAATATTTTCAGGCTATCAATCTGATATTTGATACACATA 754
234 snARGIleARgsPluDeu 240
|||||.....|||
755 ATAGAATTAAGGATCAATTA 774

seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-435-605-1

seq_documentation_block:
  Sequence 1, Application US/08435605
  GENERAL INFORMATION:
    APPLICANT: Bunette, W. Neal
    APPLICANT: Kaslow, Harvey R.
    TITLE OF INVENTION: Recombinant DNA-derived
    TITLE OF INVENTION: Cholera Toxin
    NUMBER OF SEQUENCES: 32
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Amgen Inc.
      STREET: Amgen Center
      STREET: 1840 Denavilland Drive
      CITY: Thousand Oaks
      STATE: California
      COUNTRY: USA
      ZIP: 91320-1789
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
      COMPUTER: Apple Macintosh
      OPERATING SYSTEM: Macintosh OS 7.0.
      SOFTWARE: Microsoft Word Version 5.0
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/435, 605
        FILING DATE: 05-MAY-1995
        CLASSIFICATION: 435
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 07/694,733
          FILING DATE: 02-MAY-1991
        INFORMATION FOR SEQ ID NO: 1:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 777 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double stranded
            TOPOLOGY: circular
          US-08-435-605-1

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent similarity: 94.167  Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-435-605-1 ..

Align seg 1/1 to: US-08-435-605-1 from: 1 to: 777

1 ASnGIAsPArGLeuTYrArGAlAsPSeArTPrOPrAsPglInIeTy 17
|||.....|||
55 AATGATGATTAAGTATATATCGGCGAGATTCTAGACCTCTGATGAATAAA 104
17 sArGSeRGIcLyIeUeMeTPrOArGIGlyHIsAsnGIuTYrPhesArGc 34
|||.....|||
105 GCAGTCAGGTGCTTATGAGCCAAAGAGACAGAGATGATCTTGACCCGAG 154
34 IYThGImeTAsnIlEAsnIeUyTAsPShIsAlArGyIYThGInIthr 50
|||||.....|||
155 GTACTCAATGATTAATTCACCTTTATGTCATGCGAAGAGACTCAGACG 204
|||||.....|||
51 GlyPheValArGTYrAsPAsPglTYrValSerThrSerIeUeAr 67
|||||.....|||

```



```

205 GGATTTGTAGGCACGATGATGATATGTTCCACCTCAATTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 AAGGCCCACTTAGTGGTCAAACTATATGTCGTCATCTCACTTATTT
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATATATATGTTATAGCCACTGCACGCCAACATGTTTAACTTAATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTAGGGGCATACAGTCTCTCATCCAGATGAGAACAGAACTTCTGCTTAGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrValAsnAspGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGGGATTCCTACCTCCCAATATATGATGATGATGATGATGATGATGATG 454
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGCCTGATGACAACTTACATCCTATATAGGGGCTACAGATAGATATTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGTAACTTATGATATGCTCCAGCAGCAGATGTTATGATGATGATGATG 554
167 eProProAspHisGlnAlaTrrParGluGluProTrrPleHisHisAla 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCCGCCGACCATAGACCTTGAGGGAAGACCCGTGATTCATCATCATCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CCGCGGCTTGTGGATGCTCCAGATCATCATGATGATGATGATGATGATG 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAAAGAACCCAAAGTCTAGGTAAATTCCTTGCAGATACCAATCAATGAA 704
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 AGTTAAAGACAAATATTTTCAGGCTTCAATCATCATGATGATGATGATG 754
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ATAGATTAAGATGAATTA 774
seq_name: /cgn2_6/plodata/2/pna/US094_COMB.seq:US-09-470-124-45
seq_documentation_block:
: Sequence 45, Application US/09470124
: GENERAL INFORMATION:
: APPLICANT: Mason
: TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
: FILE REFERENCE: 4866/84454
: CURRENT APPLICATION NUMBER: US/09/470,124
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 45
: LENGTH: 777
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(777)
: OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
: OTHER INFORMATION: cholera toxin gene mutagenized to optimize
: US-09-470-124-45

```

```

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent Similarity: 94.167      Percent Identity: 81.667
alignment_block:
US-09-528-682-1 x US-09-470-124-45  ..
Align seq 1/1 to: US-09-470-124-45 from: 1 to: 777
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 AATGATGACAAAGCTCTATAGGGCAGACTCAAGACCTCTGATGATGATCA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 GCAATCAGGCTGCTCTTATGCCAAGGGACAATCTGACTACTTGTACAGG 154
34 IyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GTACTCAGATGACATCAACCTTATATGACCATGCAAGGGACAATCAACT 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GGATTTGTAGGCAATGATGATGATGATGATGATGATGATGATGATGATG 254
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 GTCTGCCCACTTGGTGGTCAAACTATCTCTGTGTCACCTCACTACTACT 304
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ACATCTATGATGATGCCACGACCCCAACATGTTCAATGTAATGATGTG 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTGGGACATACACCCCTCAGCAGATGAGCAGAGAGGTGTCTGCTTTGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrValAsnAspGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGAATCCCATCTCCCAATCTATGATGATGATGATGATGATGATGATG 454
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGCCTGATGACACCTCATAGATAGGGGCTACAGGATGATGATGATGAT 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGCAACTTGGACATTCCTCCAGCAGCAGATGTTATGATGATGATGATG 554
167 eProProAspHisGlnAlaTrrParGluGluProTrrPleHisHisAla 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCCTCAGACGATAGGGCTTGAGGAGAGACCTTGATTCACCATCATCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CACGAGGTGTGGAATGCTCCAAAGGTCAAACAGATGACCAACACTTGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAAAGAACCCAAATCTTGGGTGTAAGTCTCTTATGATGATGATGATGAT 704
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 GGTGAAGGCAAACTTCTCAGGCTACCAATCTGACATTCACACCCACA 754
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ATAGATCAAGATGAACATC 774
seq_name: /cgn2_6/plodata/2/pna/PCFUS_COMB.seq:PCF-US97-11719-12

```


[illegible]

```

seq_name: /sgn2_6/prodata/2/pna/US086_COMB.seq:US-08-674-895-12
seq_documentation_block:
? Sequence 12, Application US/08674895
? GENERAL INFORMATION:
? APPLICANT: Biell, Jeffrey D.
? APPLICANT: Bookbinder, Louis H.
? APPLICANT: Hein, Mich B.
? TITLE OF INVENTION: IMMUNOCONTRACTION COMPOSITIONS
? TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
? NUMBER OF SEQUENCES: 19
? CORRESPONDENCE ADDRESS:

```

```

STATE: California
COUNTRY: United States
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,895
FILING DATE: 03-JUN-1996
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 529.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-674-895-12

```



```

Sequence, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
APPLICANT: ARLINGTON, Joshua
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: AP#41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5500
TYPE: DNA
ORGANISM: pBV2002 plasmid
US-09-724-315-1

alignment_scores:
Quality: 1088.00      Length: 240
Ratio: 4.814          Gaps: 0
Percent Similarity: 94.167      Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-09-724-315-1 ..

Align seg 1/1 to: US-09-724-315-1 from: 1 to: 5500

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
3975 AAGGATGATAGATGATATATCGGGCAGATTCCTGACCTCGATGAAATGAA 4024
||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
17 sarSerGlyGlyLeuMetProArgGlyIleHisAsnGluTyrPheAspArg 34
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4025 GCAGTCAGAGGCTCTTATGCCAAGAGCAGAGTACTTTGACCGAG 4074
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
34 IyhrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4075 GTCCTCAAGATGAATATCAACCTTATATATCATGCAAGAGCAATCAGACG 4124
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeuArg 67
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4125 GGATTTGTTAGGCACGATGATGATGTTTCCACCTCAATTAAGTTTGA 4174
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4175 AAGTGGCCCATAGTATGATGGCTCAAACTATATTTGTCGTATCTACTATTT 4224
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4225 ATATATAATGTTATAGCCACGTCACCAACATGTTAAGCTTAATGATGTA 4274
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
101 LeuGlyIleTyrSerProHisIleProTyrGluGlnGluValSerAlaLeuG 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4275 TTAGGGGCAATACACTCTCTCATCCAGATGACACAGAAAGTTTCTGCTTAG 4324
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyV 134
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4325 TGGGATTCATACATCCCAATATATATGATGGTATACGATTCATTTTGGGG 4374
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4375 TGCATTGATGAACATTTACATCGTAAATAGGGGCTACAGAGATAGATATAC 4424
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4425 AGTAACTTAATATATTTGCTCCAGCAGCAATGATGTTATGCAATTTGGCGTTT 4474
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
167 eProProAspHisGlnAlaTyrArgGluGluProTyrPheHisHisAlaP 184
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4475 CCTCCCGGACATGAGACTTGAAGGAAAGAACCCGTGATTCATCATATGCAC 4524
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysasn 200
|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```



```

4525 CCGCGGTGGGATGCTCCAAATCATCATGATGATTAATCTTCCGAT 4574
201 GUGUthrglnasnleuserthriletyrleuarglutyrglnserly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4575 GAAAAAACCAAGTAGGTAAATTCCTTACGATACCAATCAATTA 4624
217 sVallysarGlnllepheserapyrglnserGluValaspilletyrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4625 AGTTAAAGACAAATATTTTCAGCTATCAATCATGTGATGATACACATA 4674
234 snArgillearGaspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4675 ATGAAATTAAGATGAATTA 4694

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:us-08-670-974-6

seq_documentation_block:
; Sequence 6, Application US/08670974
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; APPLICANT: Waldor, Matthew K.
; TITLE OF INVENTION: NOVEL CHOLERA VECTORS, VACCINES,
; TITLE OF INVENTION: AND METHODS FOR ANTIGEN DELIVERY IN GRAM-NEGATIVE
; TITLE OF INVENTION: BACTERIA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,974
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00742/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6943 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; US-08-670-974-6

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-670-974-6 ..
Align seq 1/1 to: US-08-670-974-6 from: 1 to: 6943
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProbspGluLeu 17

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5745 AATGATGATTAAGTTATATCGGAGATTTCTAGACCTCTCATGAAATAAA 5794
17 sarGserGlyGlyleuMetProArgGlyHisasnGluTyrPhasparG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5795 GCAGTCAGTGTGCTTATGCCAAGAGCAGAGTGAATCTTGTGACCGAG 5844
34 lYthrgImetasnleasnleuTyrAspHisAlaargGlythrglnthr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5845 GTACTCAATGAATATCAACCTTTATGATCATGCAAGAGCACTCAGAGC 5894
51 GlyPheValArgTyrAspAspGlyTyrValserThrSerleuSerleuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5895 GGATTTGTTAGGCACGATGATGATGTTTCCACCCTCAATATGATTGAG 5944
67 gSerAlaHisleuAlaGlnserIleleuSerGlyTyrSerThrTyrTr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5945 AAGTGCCCACTTAGTGGGTCAAACTATATTGCTGCTGATCTTACTTATT 5994
84 yIleTyrValIleAlaThrAlaProAsnMetPheasnValaAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5995 ATATATATGTTATAGCCACGACCAACATGTTTAACGTTATATGATGA 6044
101 LeuGlyValTyrSerProHisProTyrGlnGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6045 TTAGGGGCATACAGTCTCATCAGATGACAAAGATTTCTCTTAGG 6094
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValaAspGlyV 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6095 TGGGATTCCTACTCCCAATATATGATGATGATCCAGTTTCATTTGGG 6144
134 AlileaspGluArgLeuHisArgasnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6145 TCGTGTATGAACATTTATCATCGTAAATAGGGGCTACAGAGATATGATTAC 6194
151 ArgAsnleuasnIleAlaProAlaGluAspGlyTyrArgleuAlaGlyPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6195 AGTAACTTAAATATGCTCCAGCAGCAGAGGTATGATGATGGCAGTTT 6244
167 eProProAspHisGlnAlaThrArgGlnGluProTyrPheHisAlaLap 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6245 CCCTCCGAGCATAGAGCTGGAGGAGAGCCGTGATTCATCATGCGAC 6294
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrglyAspThrCysasn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6295 CCGCGGTGTGGGAATGCTCAAGATCATCATGATTAATCTTCCGAT 6344
201 GUGUthrglnasnleuserthriletyrleuarglutyrglnserly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6345 GAAAAAACCAAGTCTAGGTAAATTCCTTACGATACCAATCAATCTTA 6394
217 sVallysarGlnllepheserapyrglnserGluValaspilletyrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6395 AGTTAAAGACAAATATTTTCAGGCTATCAATCATGTGATGATGACACATA 6444
234 snArgillearGaspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6445 ATGAAATTAAGATGAATTA 6464

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-49

seq_documentation_block:
; Sequence 49, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49

```



```

; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-49

```

```

alignment_scores:
    Quality: 1087.00      Length: 240
    Ratio: 4.831          Gaps: 0
    Percent Similarity: 93.750      Percent Identity: 81.667

```

```

alignment_block:
US-09-528-682-1 x PCT-US99-30747-49 ..

```

```

Align seg 1/1 to: PCT-US99-30747-49 from: 1 to: 777

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIlely 17
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
55 AATGATGACAAAGCTATAGGAGAGCTCAAGACCTCTCATGAGATCA 104
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
105 GCATATCAGGAGGCTTATGCCAAGGGGACAACTGAGTACTTGCACGG 154
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
155 GTCTCTCATGATGACATCAACCTTTATGACATGACAGGGAGCACTCA 204
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
205 GGATTTGTGAGGATGATGATGATGATGATGATGATGATGATGATG 254
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
255 GTCTGCCCACTTGTAGGGGTCACCAATCTCTCTGTCTGTCTACTACT 304
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
84 YrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
305 ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATG 354
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
355 TTGGGAGCATACAGCCCTCACCAGATGACAGAGAGGTCTCTCTTGG 404
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrValAsnPheGly 134
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
405 TGGAAATCCCTACTCTCCCAATCTATGATGATGATGATGATGATG 454
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
455 TCGTTGTGAGCAACTCATAGGAATAGGGGCTACAGGGATAGATAC 504
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
505 AGCAACTTGGACATGTGCTCCAGCAGCAGATGTTATGATGATGATG 554
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
167 eProProAspHisGlnAlaTyrPargIleGluIleProTyrPheHis 184
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
555 CCTCTGAGCAATAGGGCTTGGAGGGGAGGCTTGGATTCACCATG 604
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
605 CACCAAGTTTGGAAATGCTCCAAAGGTCAAGCATGACCAACACTGT 654
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```

655 GAAAGACCCCAATCTTGGGTGTAAGTCTCTGATGATGATCAATCTAA 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGlnValAlaPheI 234
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
705 GGTGAAGAGGCAATCTTCTCAGGCTACCAATCTGACATTGACACCA 754
234 snArgIleArgAspGluLeu 240
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
755 ATAGATCAAGATGATGAATCTC 774

```

```

seq_name: /cgn2_6/plodata/2/pna/US094.COMB.seq:US-09-470-124-49

```

```

seq_documentation_block:
; Sequence 49, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arntzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-49

```

```

alignment_scores:
    Quality: 1087.00      Length: 240
    Ratio: 4.831          Gaps: 0
    Percent Similarity: 93.750      Percent Identity: 81.667

```

```

alignment_block:
US-09-528-682-1 x US-09-470-124-49 ..

```

```

Align seg 1/1 to: US-09-470-124-49 from: 1 to: 777

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIlely 17
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
55 AATGATGACAAAGCTATAGGAGAGCTCAAGACCTCTCATGAGATCA 104
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
55 AATGATGACAAAGCTATAGGAGAGCTCAAGACCTCTCATGAGATCA 104
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
105 GCATATCAGGAGGCTTATGCCAAGGGGACAACTGAGTACTTGCACGG 154
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
155 GTACTCAGATGACATCAACCTTTATGACCATGCAAGGGAGCACTCA 204
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
205 GGATTTGTGAGGATGATGATGATGATGATGATGATGATGATGATG 254
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
255 GTCTGCCCACTTGTAGGGGTCACCAATCTCTCTGTCTGTCTACTACT 304
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
84 YrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
305 ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATG 354
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```



```

; SEQ ID NO 47
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
; US-09-470-124-47

```

```

alignment_scores:
  Quality: 1084.00      Length: 240
  Ratio: 4.796         Gaps: 0
  Percent Similarity: 94.167   Percent Identity: 81.250

```

```

alignment_block:
  US-09-528-682-1 x US-09-470-124-47

```

```

Align seg 1/1 to: US-09-470-124-47 from: 1 to: 777

```

```

1  AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
55  AATGATGACAGGCTCTATAGGCGAGACTCAAGACCTCTGATGATGATCA 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17  sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTYrPheAspArg 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105  GCATATCAGGTGGTCTTATGCCAAGGGGAGCAATCTGACTATTGACAGG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34  lyThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155  GTACTCAGATGACATCAACCTTTATACCATGCAAGGGGAGACTCAACT 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51  GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205  GGATTTGTGAGGATGATGATGATGTGTGCCAAGATTTACTTGAG 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYr 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255  GTCTGCCCATTTGGTGGTCAAACTATCTCTGTCACCTCTACTACT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84  yTleTYrValIleAlaThrAlaProAsnMetPheAsnValAspAspVal 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305  ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101  LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeu 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355  TTGGGAGCATACAGCCCTCAACCCAGATGACAGAGGCTGTCTGTGGG 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117  yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPheGly 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405  TGGATATCCCACTACTCCCAATCTATGATGATGATGATGATGATG 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134  alIleAspGluArgLeuHisArgAsnArgGluTYrArgValAspArgTYr 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455  TGGCTGTGAGCATCTCCATAGGAGGCTACAGGAGATGATGATGATG 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505  AGCAATTTGGCATTTGCTCCAGCAGCATGTTATGATGATGATGATG 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167  eProProAspHisGlnAlaTrpArgGluGluProTYrIleHisAlaP 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
555  CCCGCCAGACATAGGCTTGGAGGAGGAGCCCTTGATTCACCATG 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184  roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
605  CACCAAGTTGTGAAATGCTCCAAAGTCAAGCATGACCAACTTGTG 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201  GlnGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSer 217

```

```

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
655  GAAAGACCCAACTTTGGGTGAGAGTCTCTGATGATGATGATGATG 704
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217  sVallyArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
705  GGTGAAGAGGCAAACTTCTCAGGCTACCAATCTGACATTGACACCC 754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234  snArgIleArgAspGluLeu 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
755  ATAGATCAAGATGAATC 774

```

```

seq_name: /cgn2_6/plodata/2/pna/pctus.COMB.seq: PCT-US99-30747-51

```

```

seq_documentation_block:
  Sequence 51, Application PC/TUS9930747

```

```

; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-51

```

```

alignment_scores:
  Quality: 1081.00      Length: 240
  Ratio: 4.804         Gaps: 0
  Percent Similarity: 93.750   Percent Identity: 81.250

```

```

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-51

```

```

Align seg 1/1 to: PCT-US99-30747-51 from: 1 to: 777

```

```

1  AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
55  AATGATGACAGGCTCTATAGGCGAGACTCAAGACCTCTGATGATGATCA 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17  sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTYrPheAspArg 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105  GCATATCAGGTGGTCTTATGCCAAGGGGAGCAATCTGACTATTGACAGG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34  lyThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155  GTACTCAGATGACATCAACCTTTATACCATGCAAGGGGAGACTCAACT 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51  GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205  GGATTTGTGAGGATGATGATGATGATGTGTGCCAAGATTTACTTGAG 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYr 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255  GTCTGCCCATTTGGTGGTCAAACTATCTCTGTCACCTCTACTACT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84  yTleTYrValIleAlaThrAlaProAsnMetPheAsnValAspAspVal 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305  ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101  LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeu 117

```


51 01:00:00


```

|||||.....
722 GGATTTGGTGGCAGATGATGATGTTTCCACCTCAATTAAGTTGAG 771
67 gserAlahisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
|||||.....
772 AAGTGGCCACTTAAAGTGGTCAAACTATATGTCGTCATCTCTATT 821
84 yrlIeYValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||.....
822 ATATATATGTTATAGCCACCTCACCCAACTGTTTAACTTAATGATGA 870
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||.....
871 TTAGGGGCAATACAGTCCATCCATCCAGATGAACAAGATTCCTGCTTAG 920
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhelGly 134
|||||.....
921 TGGGATTCACATCTCCCAATATATGATGATGATTCGATCTTTGGGG 970
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....
971 TCGTTGATGAACAAATTAATCATGTAATAGGGGCTACAGAGATAGATTTAC 1020
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||.....
1021 AGTAACTTAAGTATGCTCCAGACAGATGTTATGATGATGGCAGGTT 1070
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAlaP 184
|||||.....
1071 CCCGCCGAGATAGACCTTGAGAGGAAGACCCCTGATTCATCATGCAC 1120
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||.....
1121 CGCCGGGATGGGGAATGCTCCAAAGATCATGATGATTAATCTTCGAT 1170
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerI 217
|||||.....
1171 GAAAAAACCCAAAGCTGAGGTGTAATAATCCTTACGAAATACCAATCTAA 1220
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||.....
1221 AGTTAAAGAACAAATATTTCAGGCTATCATGATATGATACACATA 1270
234 snArgIleArgAspGluLeu 240
|||||.....
1271 ATAGAAATTAAGATGAATTA 1290

```

seq_name: /cgn2_6/prodata/2/pna/PCrTUS_COMB.seq:PCr-US01-08582-2

```

seq_documentation_block:
; Sequence 2, Application PC/TUS0108582
; GENERAL INFORMATION:
; APPLICANT: UAB Research Foundation
; TITLE OF INVENTION: Chimeric Nontoxic Mutants of
; Humoral Immunity
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc
; STREET: P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US01/08582
; FILING DATE: 16-Mar-2001
; CLASSIFICATION: <Unknown>

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: MCG-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/425-8405
; TELEFAX: 703/425-8406
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2022 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
PCT-US01-08582-2

```

```

alignment_scores:
Quality: 1049.00 Length: 240
Ratio: 4.683 Gaps: 0
Percent Similarity: 93.33 Percent Identity: 80.417

```

alignment_block:

US-09-528-682-1 x PCT-US01-08582-2

Align seg 1/1 to: PCT-US01-08582-2 from: 1 to: 2022

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
||| |||.....
572 AATGATGATGAATTAATTCGCGGCGAGATTCAGACCTCGATGAATAAA 621
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||.....
622 GCACTCAGGTGGTCTTATGCCAAGACCACAGAGAGTACTTGGACCGAG 671
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||.....
672 GTACTCAATAGAAATATACACCTTATGATCATGCAGAGAGACACAGAG 721
672 GATTTGTTAGGCCAGATGATGATGTTTCCACCTCAATTAAGTTTGGAG 771
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||.....
722 GGATTTGTTAGGCCAGATGATGATGTTTCCACCTCAATTAAGTTTGGAG 771
67 gserAlahisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyr 84
|||||.....
772 AAGTGGCCACTTAAAGTGGTCAAACTATATGTCGTCATCTCTATT 821
84 yrlIeYValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||.....
822 ATATATATGTTATAGCCACCTCACCCAACTGTTTAACTTAATGATGA 870
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||.....
871 TTAGGGGCAATACAGTCCATCCATCCAGATGAACAAGATTCCTGCTTAG 920
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhelGly 134
|||||.....
921 TGGGATTCACATCTCCCAATATATGATGATGATTCGATCTTTGGGG 970
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....
971 TCGTTGATGAACAAATTAATCATGTAATAGGGGCTACAGAGATAGATTTAC 1020
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||.....
1021 AGTAACTTAAGTATGCTCCAGACAGATGTTATGATGATGGCAGGTT 1070
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAlaP 184
|||||.....
1071 CCCGCCGAGATAGACCTTGAGAGGAAGACCCCTGATTCATCATGCAC 1120

```



```

36  lmetasnleasnleutyaspHisalAargglyThrGlnThrGlyPhe 52
|||||
109  AATGATATCAACCTTATGATCATGCAGAGAACTCAGCGGATTT 158
|||||
53  ValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAl 69
|||||
159  GTTAGGACGATGATGATGATGTTCCACCTCAATTAGTTGAGAAAGTGC 208
|||||
69  aHlsleuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTrpIleT 86
|||||
209  CCACCTTAGTGGTCAACCTATATTGCTGTGTCATTTCTACTTATTATATAT 258
|||||
86  yValAlIleAlaTrpAlaProAsnMetPheAsnValAsnAspValLeuGly 102
|||||
259  ATGTTATAGCACCTGCACCCACACATGTTTAACTTATGATGATTTAGG 308
|||||
103  ValTyrSerProHisProTyrGlnGlnValSerAlaLeuGlyIle 119
|||||
309  GCATACAGTCTCATCCAGATGAACAAGAGTTTCTCTTAGCGTGGAT 358
|||||
119  eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhgGlyValIleAl 136
|||||
359  TCCATACTCCCAATATATGATGATGATGATTCATTTTGGGCTGTG 408
|||||
136  spGIuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
409  ATGACACATTACATCTAATAGGGGCTACAGACATGATATATACAGTAAC 458
|||||
153  LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
459  TTGATATCTCTCCAGCAGCAGATGATTGATGATGATGATGATGATGATG 508
|||||
169  oAspHisGlnAlaTrpArgGlnGluProTyrIleHisIleAlaProGln 186
|||||
509  GGAGCATAGAGCTTGGAGGAGAGCCGTGATTCATCATGACGCCGCGG 558
|||||
186  lYcysGlyAsnSerSerArgThr 193
|||||
559  GTTGTGGAGATGCTCCAAAGATCA 581
|||||
seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-402-100-1

seq_documentation_block:
; Sequence 1, Application US/09402100
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter pylori
; FILE REFERENCE: 0136/06140
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

```

```

alignment_scores:
  Quality: 252.00      Length: 112
  Ratio: 3.111        Gaps: 3
  Percent Similarity: 72.321      Percent Identity: 47.321

alignment_block:
US-09-528-682-1 x US-09-402-100-1 ..

Align seg 1/1 to: US-09-402-100-1 from: 1 to: 2385

137  GluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArg.... 151
|||||
1557  CAGAAATATCCTAAAGAAAGACATGCAATTCACAGACACACCGCTCACAA 1606
|||||
152  .....AsnLeuAsnIleAlaProAlaGluAspGlyTyrAla 163
|||||
1607  TTGAAGTCAATTCTGAACCTTA.....CCATGCTTCGTGATGGC 1647
|||||
163  rGluAlaGlyPheProProAspHisGlnAla.TrpArgGlnGluProTyr 179
|||||
1648  AAGACATACCTCTAAACCGCAATTAAGTGAATTCGAAAGACCGCTG 1697
|||||
179  rIleHisIleAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrG 196
|||||
1698  GATTCATCATGACACCGCGGCTGTGGAATGCTCCAAAGATCATCATCA 1747
|||||
196  lYsAspThrCysAsnGlnGluTyrGlnAsnLeuSerThrIleTyrLeuArg 212
|||||
1748  GTATACTCTTCGATGAAAAAACCMAAGTGTAGGTGAATAAATTCCTTGAC 1797
|||||
213  GluTyrGlnSerYsValYsArgGlnIlePheSerAspTyrGlnSerG 229
|||||
1798  GAATACCAATCTAAGTTAAAGACAAATATTTTTCAGGCTATCATCACTGA 1847
|||||
229  vAlaAspIleTyrAsnArgIleArgAspGluLeu 240
|||||
1848  TATTGATACACATTAATTAAGATGATTA 1881
|||||

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-402-100-3

seq_documentation_block:
; Sequence 3, Application US/09402100
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter pylori
; FILE REFERENCE: 0136/06140
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4149
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-3

alignment_scores:

```


Quality: 245.00 Length: 65
Ratio: 4.298 Gaps: 0
Percent Similarity: 87.692 Percent Identity: 66.154

alignment_block:

US-09-528-682-1 x US-09-402-100-3 ..

Align seg 1/1 to: US-09-402-100-3 from: 1 to: 4149

```

176 GUGUURPOTPIEHSHSALAPROGLNGLYSGLYSNERSEAR 192
|||||
3451 GAGAGCGCGATTCAATGACGCCCGGTGTGGAACTCTCCAG 3500
|||||
192 gthclthrclyasphrcysasnglugluthrclnleuserrhri 209
|||||
3501 ATCTTCGATCAGTAACTCTGCGATGAAAAACCAAGTCTAGGTGTA 3550
|||||
209 leryleuargglutrglnserlysalysarglnlepheserasp 225
|||||
3551 AATTCCTTGAGATACCAATCTAAAGTTAAAGCAAAATTTTCAGGC 3600
|||||
226 TyrcInsergluValaspIleTyraNargIleargaspIuleu 240
|||||
3601 TATCAATCTGATATGATACACATATAGATTAAGATGATTA 3645
|||||
seq_name: /cgn2_6/prodata/2/pna/US090_COMB.seq:US-09-051-315-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09051315
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co., Ltd
; APPLICANT: Kim, Byung-O
; APPLICANT: Lee, Byung-Kwang
; APPLICANT: Ioon, Suk-Won
; APPLICANT: Park, Seung-Kook
; APPLICANT: Yu, Young-Hyo
; TITLE OF INVENTION: A RECOMBINANT MICROORGANISM EXPRESSING
; TITLE OF INVENTION: AN ANTIGENIC PROTEIN, ADHESIN
; FILE REFERENCE: 0136/06164
; CURRENT APPLICATION NUMBER: US/09/051,315
; EARLIER FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: PCT/KR 97/00091
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of a fusion gene prepared by ligating adhesin gene c
; OTHER INFORMATION: A2 and B subunit genes of Vibrio cholerae toxin.
; US-09-051-315-1

```

alignment_scores:

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

alignment_block:

US-09-528-682-1 x US-09-051-315-1 ..

Align seg 1/1 to: US-09-051-315-1 from: 1 to: 1516

```

1 AsnGlyAspArgLeuTyraGAlaAspSerArgProProaspGluIlely 17
|||||
403 AATGCGCAATGTGTTACGCCGATCTAAAGACCAATACAGAAAAA 452
|||||
17 SARSerIlyglIleuMet..... 23
|||||
453 ATCAGAACCGCGTTATATCTCCACCGGTTTGACAAAAATGGAAGCG 502
|||||

```

```

24 .....ProArgLy.....His 27
|||||
503 TTTTAATCCCGCGTGGTTTATTAAAGTTACCATCTAGACCTATGAGT 552
|||||
28 AsnGluTyrrheaspargglyThrGlnmecasnlleasnleuTyraSpH 44
|||||
553 GCGGAATCTTGGATCTTTTACGATGATTTGAGCGGTTGGACATATCA 602
|||||
44 salargGlythrGlnThrGlypheValArgTyraSpasglyTyraVal 61
|||||
603 AGAAATAATCTTAAACACCATTCACCATACGCGGGGTTAGTTA 652
|||||
61 erThSerleuSerleuArgSerleuAhlsleuAlagIlnserIleu 77
|||||
653 GCACCT.....ATGTTAAGGAGAACGATATCTTAATGACGCATCAG 696
|||||
78 SerGlyTyrrSerThrTyrrIleTyraValIlealThrAlaProasme 94
|||||
697 AGCGCTTGAATAGATTTTGCATAATATC.....AT 728
|||||
94 tPheasnValasnpaspyalleuGlyValTyrrSerProHisProtyrglu 111
|||||
729 GCAAGAAATAGACAAAAACCTCACTCAAAAGAAATTAAGAACTTATCAAA 778
|||||
111 lngluValserAlaleuGlyIleProtyrSerGlnIleTyrglyTrr 127
|||||
779 AAGACGCCAAAGATTTAAAGGC..... 801
|||||
128 TyrrArgValasnpheGlyValIleaspGluArgleuHisArgasnarGcl 144
|||||
802 .....AAAAAACCAGCA 815
|||||
144 uTyrrArgspargTyrrTyrrArgasnlleuasnIlealThrAlaGluasp 161
|||||
816 ATTC..... 819
|||||
161 lyTyrrArgleuAlaglyPheProProaspHisGlnAlatrPargGlu 177
|||||
820 .....CAG 822
|||||
178 ProTrrIleHSHSALAPROGLNGLYSGLYSNERSEARgthri 194
|||||
823 CCGTGATTCATCATGACCCCGGTGTGGAAATGCTCAAGATCATC 872
|||||
194 ethrclyasphrcysasnglugluthrclnleuserrhriIleTyrl 211
|||||
873 GATCAGTAACTCTGCGATGAAAAAACCAAGCTAGTAAATTC 922
|||||
211 euArgGluTyrglnserlysalysarglnlepheserAspTyrgln 227
|||||
923 TTGACGAATACCAATCTAAAGTTAAAGACAAATATTTTCAGGCTATCAA 972
|||||
228 SerGluValaspIleTyraNargIleargaspIuleu 240
|||||
973 TCTGATATTGATACACATATAGATTAAGATGATTA 1011
|||||

```

seq_name: /cgn2_6/prodata/2/pna/US094_COMB.seq:US-09-423-493-1

seq_documentation_block:

```

; Sequence 1, Application US/09423493
; GENERAL INFORMATION:
; APPLICANT: DAEMONG PHARMACEUTICAL CO., LTD
; APPLICANT: Kim, B.O., et al.
; TITLE OF INVENTION: A preventive and therapeutic vaccine for
; TITLE OF INVENTION: Helicobacter pylori-associated diseases.
; FILE REFERENCE: 0136/06320
; CURRENT APPLICATION NUMBER: US/09/423,493
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: PCT/KR98/00072
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```



```

; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion oligonucleotide between Helicobacter pylori
; OTHER INFORMATION: and Vibrio cholerae
US-09-423-493-1

```

alignment_scores:

```

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

```

alignment_block:

```

US-09-528-682-1 x US-09-423-493-1 ..

```

```

Align seg 1/1 to: US-09-423-493-1 from: 1 to: 1516

```

```

1 AsnGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIleTy 17
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
403 AATGCCGAATGTGTTTACGCCCGATCCTAAAGAACATACAGAAAAA 452
17 sArgSerGlyGlyLeuMet..... 23
453 ATCAGAACCCGGGTATATTCCTCCACCGGTTTGACAAATGGAAGGGG 502
24 .....ProArgGly.....His 27
503 TTTTAATCCCGCGGTGTTTAAAGTTACCATACAGCCATAGT 552
28 AsnGluTyRpheAspArgGlyThrGlnMetAsnIleAsnLeuTyRAspH 44
: : : : : : : : : : : : : : : : : : : : : : : : : : :
553 GGGGAATCTTGGATCTTTTACGATGATTTGACGAGTTGACATTC 602
44 sAlaArgGlyThrGlnThrGlyPheValArgTyRAspAspGlyTyRVal 61
: : : : : : : : : : : : : : : : : : : : : : : : : : :
603 AGAAAAATCTTAAACACCATTCACATGACGCGGGGTAGCTTA 652
61 eThrSerLeuSerLeuArgSerAlaHisLeuAlaGlnSerIleLeu 77
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
653 GCACT.....ATGGTTAAGGAACGATATTTCTAATGACGATCAAG 696
78 SerGlyTyRSerThrTyRLeuTyRValIleAlaThrAlaProAsnMe 94
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
697 AGCCCTTGTGATTAAGATTTTGCATAATATC.....AT 728
94 tPheAsnValAsnAspValLeuGlyValTyRSerProHisProTyRGl 111
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
729 GCAAGAAATGACAAAAAACTCCTCAAAAGAAATTTAGAAATCTTATC 778
111 lngIuValSerAlaLeuGlyGlyIleProTyRSerGlnIleTyRGl 127
: : : : : : : : : : : : : : : : : : : : : : : : : : :
779 AAGACGCCAAGATTTAAAGC..... 801
128 TyRArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArg 144
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
802 .....AAAAGAAACGAGA 815
144 uTyRArgAspArgTyRtyrArgAsnLeuAsnIleAlaProAlaGlu 161
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
816 ATTC..... 819
161 lTyRArgLeuAlaGlyPheProProAspHisGlnAlaTrpArgGlu 177
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
820 .....GAG 822
178 ProTPrIleHisAlaProGlnGlyCysGlyAsnSerSerArgTrpH 194
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
823 CCGTGATTCATCATGACCGCGGTTTGGAATCTCCACATCATC 872
194 eThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyR 211
: : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

873 GATCAGTAATACCTTCGATGATGAAAAAACCCAAAGTCTAGTCTAAATTC 922
211 eUArgGluTyRGlSerLeuValysArgGlnIlePheSerAspTyRGl 227
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
923 TTGACGAATTCACATCTTAAAGTTAAAGAACAAATATTTTACGGCTAT 972
228 SerGluValAspIleTyRAsnArgIleArgAspGluLeu 240
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
973 TCTGATATTCATACATATTAATAGAAATTAACGATGAATTA 1011
seq_name: /cgn2.6/plodata/2/pna/US097C_COMB.seq:US-09-760-234-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09760234
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co., Ltd
; APPLICANT: Kim, Byung-O
; APPLICANT: Lee, Byung-Kwang
; APPLICANT: Yoon, Suk-Won
; APPLICANT: Park, Seung-Kook
; APPLICANT: Yu, Young-Hyo
; TITLE OF INVENTION: A RECOMBINANT MICROORGANISM EXPRESSING
; TITLE OF INVENTION: AN ANTIGENIC PROTEIN, ADHESIN
; FILE REFERENCE: 0136/IE16-US1
; CURRENT APPLICATION NUMBER: US/09/760,234
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: PCT/KR 97/00091
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of a fusion gene prepared by ligating
; OTHER INFORMATION: adhesin gene of H. pylori and A2 and B subunit genes of
; OTHER INFORMATION: Vibrio cholerae toxin
US-09-760-234-1

```

alignment_scores:

```

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

```

alignment_block:

```

US-09-528-682-1 x US-09-760-234-1 ..

```

```

Align seg 1/1 to: US-09-760-234-1 from: 1 to: 1516

```

```

1 AsnGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIleTy 17
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
403 AATGCCGAATGTGTTTACGCCCGATCCTAAAGAACATACAGAAAAA 452
17 sArgSerGlyGlyLeuMet..... 23
453 ATCAGAACCCGGGTATATTCCTCCACCGGTTTGACAAATGGAAGGGG 502
24 .....ProArgGly.....His 27
503 TTTTAATCCCGCGGTGTTTAAAGTTACCATACATAGACCTATAGT 552
28 AsnGluTyRpheAspArgGlyThrGlnMetAsnIleAsnLeuTyRAspH 44
: : : : : : : : : : : : : : : : : : : : : : : : : : :
553 GGGGAATCTTGGATCTTTTACGATGATTTGACGAGTTGACATTC 602
44 sAlaArgGlyThrGlnThrGlyPheValArgTyRAspAspGlyTyRVal 61
: : : : : : : : : : : : : : : : : : : : : : : : : : :
603 AGAAAAATCTTAAACACCATTCACATGACGCGGGGTAGCTTA 652

```



```

61 eRthSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 77
||||| : : : : : : : : : : : : : : : : : : : : : : : :
653 GCACCT.....ATGCTTAAGCAACGATAATCTAATGACGGATCAAG 696
78 SerGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMe 94
||||| : : : : : : : : : : : : : : : : : : : : : : : :
697 AGCGCTTTGAAATGAATTTTGCAGAAATATC.....AT 728
94 tPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG 111
| : : : : : : : : : : : : : : : : : : : : : : : :
729 GCAGAAATAGCAAAAAAAGCTCACTCAAAAGAAATTAGAAATCTTATCTCAA 778
111 InGluValSerAlaLeuGlyIleProTyrSerGlnIleTyrGlyTyr 127
: : : : : : : : : : : : : : : : : : : : : : : :
779 AAGACCCCAAGAAATTAAAGGC..... 801
128 TyrArgValAsnDheGlyValIleAspGluArgLeuHisArgAsnArg 144
| : : : : : : : : : : : : : : : : : : : : : : : :
144 uTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAsp 161
| : : : : : : : : : : : : : : : : : : : : : : : :
816 ATTC..... 819
161 LyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTArgGluGlu 177
| : : : : : : : : : : : : : : : : : : : : : : : :
820 .....GAG 822
178 ProTyrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThrIle 194
||||| : : : : : : : : : : : : : : : : : : : : : : : :
823 CCGTGATTCATCATGACCGCCGGTGGGAATGCTCCAAAGATCATC 872
194 eThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 211
| : : : : : : : : : : : : : : : : : : : : : : : :
873 GATCAGTAATACTTGCATGAAAAAACCCAAAGCTAGGTAAATATTC 922
211 euArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
| : : : : : : : : : : : : : : : : : : : : : : : :
923 TTGACGAATTCATCTAAAGTTAAAGCAAAATATTTTACGGCTATCAAA 972
228 SerGluValAspIleTyrAsnArgIleArgAspGluLeu 240
| : : : : : : : : : : : : : : : : : : : : : : : :
973 TCTGATATTGATACATATATAGAAATTAAGATGAATTA 1011

```

seq_name: /cgn2_6/ptodata/2/pna/US081.COMB.seq:US-08-133-438-3

seq_documentation_block:

```

; Sequence 3, Application US/08133438
; GENERAL INFORMATION:
; APPLICANT: KAPER, James B.
; APPLICANT: LEVINE, Myron M.
; TITLE OF INVENTION: VIBRIO CHOLERAE NON-01 SEROGROUP VACCINE
; TITLE OF INVENTION: STRAINS, METHODS OF MAKING SAME AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & McKenzie
; STREET: Suite 1100, 815 Connecticut Ave., N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20006-4078
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,438
; FILING DATE: 08-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/472,276
; FILING DATE: 04-MAR-1983

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/581,406
; FILING DATE: 17-FEB-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/867,633
; FILING DATE: 27-MAY-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/363,383
; FILING DATE: 05-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,315
; FILING DATE: 05-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/821,872
; FILING DATE: 16-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/931,943
; FILING DATE: 12-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kile, Bradford E.
; REGISTRATION NUMBER: 25,223
; REFERENCE/DOCKET NUMBER: BAMCZ 0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 452-7000
; TELEFAX: (202) 452-7074
; TELEX: 89552
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7076 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Vibrio cholerae
; STRAIN: CVD110
; US-08-133-438-3

```

alignment_scores:

Quality	Ratio	Length	Gaps
160.50	3.567	50	1

Percent Similarity: 90.000 Percent Identity: 64.000

alignment_block:

```

US-09-528-682-1 x US-08-133-438-3/rev ..
Align seg 1/1 to reverse of: US-08-133-438-3 from: 1 to: 7076
191 SerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeu 207
||||| : : : : : : : : : : : : : : : : : : : : : : : :
1339 TCTAAGCAGATCAGT...AATACTTGCATGAAAAAACCCAAAGTCTAAG 1293
207 rThrIleTyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePhe 224
: : : : : : : : : : : : : : : : : : : : : : : :
1292 TGTAAATTCCTTGACGAATACCATCTAAAGTTAAAGCAAAATATTTT 1243
224 eAspTyrGlnSerGluValAspIleTyrAsnArgIleArgAspGluLeu 240
| : : : : : : : : : : : : : : : : : : : : : : : :
1242 CAGCGTATCATCTGATATTGATACATATATAGAAATTAAGATGAATTA 1193

```

seq_name: /cgn2_6/ptodata/2/pna/US081.COMB.seq:US-08-133-439-3

seq_documentation_block:

```

; Sequence 3, Application US/08133439
; GENERAL INFORMATION:
; APPLICANT: KAPER, James B.
; APPLICANT: LEVINE, Myron M.
; TITLE OF INVENTION: VIBRIO CHOLERAE CVD111, METHOD OF MAKING
; TITLE OF INVENTION: SAME, AND VACCINE DERIVED THEREFROM
; NUMBER OF SEQUENCES: 3

```



```

CORRESPONDENCE ADDRESS:
ADDRESS: Baker & McKenzie
STREET: Suite 1100, 815 Connecticut Ave., N.W.
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20006-4078
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,439
FILING DATE: 08-OCT-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 06/472,276
FILING DATE: 04-MAR-1983
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 06/581,406
FILING DATE: 17-FEB-1984
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 06/867,633
FILING DATE: 27-MAY-1986
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/363,383
FILING DATE: 05-JUN-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/533,315
FILING DATE: 05-JUN-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/821,872
FILING DATE: 16-JAN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/931,943
FILING DATE: 12-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kile, Bradford E.
REGISTRATION NUMBER: 25,223
REFERENCE/DOCKET NUMBER: BANCZ 0020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 452-7000
TELEFAX: (202) 452-7074
TELEX: 89552
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 7076 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: CVD110
US-08-133-439-3

alignment_scores:
Quality: 160.50 Length: 50
Ratio: 3.567 Gaps: 1
Percent Similarity: 90.000 Percent Identity: 64.000

alignment_block:
US-09-528-682-1 x US-08-133-439-3/rev ..
Align seg 1/1 to reverse of: US-08-133-439-3 from: 1 to: 7076
191 SerArgThrIleHrgIAspHrCysAsnGluGluThrGlnAsnLeu 207
|||||:|||||: :|||||:|||||:|||||:|||||:
1339 TCTAGAGCGATCAGT...ATACTTGCATGAGAAAAACCCAAAGTCTAGG 1293

```

```

207 rThrIleTyrlLeuArgIuTyrgInSerLySvalTySArgInIlePheS 224
::: ||||| ||||||| ||||||| ||||||| |||||||
1292 TCTAAATTCCTTGACGAAATACCAATCTAAAGTTAAAGACAAATTTT 1243
1242 CAGGCTATCATCTGATATTGATACACATATAGATTAAAGATGAATTA 1193
224 eRAspTyrgInSerGluValAspIleTyrsArnArgIleArgAspGluLeu 240
|| |||||||:||||| :|||||:|||||:|||||:|||||
1242 CAGGCTATCATCTGATATTGATACACATATAGATTAAAGATGAATTA 1193

seq_name: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-771-536B-15

seq_documentation_block:
; Sequence 15, Application US/09771536B
; GENERAL INFORMATION:
; APPLICANT: Langridge, William H. R.
; APPLICANT: Yu, Jie
; APPLICANT: Arakawa, Takeshi
; TITLE OF INVENTION: Transgenic Plant-Based Vaccines
; FILE REFERENCE: 12273-3
; CURRENT APPLICATION NUMBER: US/09/771,536B
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 651
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Vibrio cholerae and Escherichia coli
US-09-771-536B-15

alignment_scores:
Quality: 158.00 Length: 63
Ratio: 3.362 Gaps: 1
Percent Similarity: 74.603 Percent Identity: 50.794

alignment_block:
US-09-528-682-1 x US-09-771-536B-15 ..
Align seg 1/1 to: US-09-771-536B-15 from: 1 to: 651
188 GlyAsnSerSerArgThrIleHr..... 195
||||| ||| :|||:
460 GGAACATATTCAGAGTAGATCTGTGAATGACTTGGATCCGTCA 509
196 ...GlyAspHrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrl 211
:|||||:|||||:|||||:|||||:|||||:|||||:
510 CATCACTAATACTTGGATGAAAAACCCAAAGCTAGGTAAATTC 559
211 euArgIuTyrgInSerLySvalTySArgInIlePheSerAspTyrgIn 227
|| ||||||| ||||||| ||||||| ||||||| |||||||
560 TTGACGAATRCACATCTTAAGTTAAAGACAAATATTTCAGGCTATCA 609
228 SerGluValAspIleTyrsArnArgIleArgAspGluLeu 240
|||||:||||| :|||||:|||||:|||||:|||||
610 TCTGATATTCATACATACATATTAAGATTAAAGATGAGCTTG 648

seq_name: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-771-536B-13

seq_documentation_block:
; Sequence 13, Application US/09771536B
; GENERAL INFORMATION:
; APPLICANT: Langridge, William H. R.
; APPLICANT: Yu, Jie
; APPLICANT: Arakawa, Takeshi
; TITLE OF INVENTION: Transgenic Plant-Based Vaccines
; FILE REFERENCE: 12273-3
; CURRENT APPLICATION NUMBER: US/09/771,536B
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13

```



```

62  ThisSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
    |||:||||| :|||  :|||
370  ACTGATCTAGCTATGCGCA.....GTTCGTGCAAGAGCTTATGGA 410
    78  rGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetP 95
    |||  :|||  |||  :|||
411  TGGATGTCCTTGTGTCTACATTCAGAAAAT.....CTTGAACAATCT 454
    95  heAsnValAsnAspValLeuGlyValTyr...SerProHisProTyrGlu 110
    ||:||||| :|||  :|||  :|||
455  TCACAAAT.....CTTTTGCAATCTCTACCTTCATACACACACATATTAAT 498
    111  GlnGlnValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyT 127
    :|||  :|||  :|||  :|||
499  CTTCCATCACCATTCTATTAATGGAGTTCGCTGGCAACAATTG...GGAG 545
    127  pTyArgValAsnPhgGlyValIleAspGluArgLeuHisArgAsnArg 144
    |  :|||  |||  :|||
546  G.....TTAATCCATGAATTCGCACAT...CATCAGT 574
    144  LuTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAsp 160
    :|||  :|||  :|||  :|||
575  TGTTCAAAAACAGATACTACTACATGATTTG.....GCCAGTATTTGCTT 618
    161  GlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTPrArgGlu 177
    |||  |||  |||  :|||
619  GGAACCTTTTTCACAGAGATTC.....TCATCTGGTGGTGGAAAAGCA 662
    177  uProTPrIleHisHisIleA...ProGlnGlyCysGlyAsnSerSerArgTln 193
    :|||  :|||  :|||  :|||
663  GCACATGTGTCATCCAGCAGCACCAAAATGTTGTGGACGACGCGGATC 712
    194  IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
    :|||  :|||  :|||  :|||
713  TTGATTATGATCCCATCTCTATGCTACAGTGGCAGACATCTCAACATTAAT 762
    210  rLeuAlaGlu..... 213
    |||
763  TCTCAGAGATTTCATGGGTTATGACTCTATTCAGATGCAACATGTTCAATTG 812
    214  .....TyrGlnSerLysValLysArgGlnIlePheSer 224
    |||:||||| :|||  :|||
813  GACATTCATGCTTACCATCTCTCCGCTCTCTGCTGCTTCTTCAGT 856

```

```

1 NAME: Sutton, Jeffrey A
2 REGISTRATION NUMBER: 34028
3 REFERENCE/DOCKET NUMBER: 12098
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: 215 270 5024
6 TELEFAX: 215 270 5090
7 INFORMATION FOR SEQ ID NO: 1:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 4207 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: single
12 TOPOLOGY: linear
13 MOLECULE TYPE: DNA (genomic)
14 HYPOTHETICAL: NO
15 ANTI-SENSE: NO
16 US-08-245-848-1
17
18 Alignment scores:
19 Quality: 96.50 Length: 217
20 Ratio: 1.016 Gaps: 10
21 Percent Similarity: 43.779 Percent Identity: 24.424
22
23 Alignment block:
24 US-09-528-682-1 x US-08-245-848-1 ..
25
26 Align seg 1/1 to: US-08-245-848-1 from: 1 to: 4207
27
28 5 LeuTYrAArgLaAspSerArgProProAspGluIleLysArgSerGly 21
29 ::::::::::|::::::::::::::::::::::::::|: 22
30 627 GATACCGCTATGACTCTCCGCCCGGAGGAGGTTTCCAGAAC..GG 673
31
32 21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
33 ::::::::::|::::::::::::| 37
34 674 ATTCACGGCGCTGGGGAACAACAGAC..... 698
35
36 38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
37 ::::::::::|::| 53
38 699 ....AATGGCTCGACCATCTGCACCGGACCTCTCCAGGTCGGCAC 743
39
40 55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArg..... 67
41 ::::::::::|::::::::| 66
42 744 AGCAACAGCGCTTTCGCTCCACGACGACGCGGCGCTATACGAGGT 793
43
44 68 .....SerAlaHisLeuAlaGly 74
45 ::::::::::| 73
46 794 CTAATCGAACATCCGATCGACGAAAGCGTTGAGGCCAACCGCGCGCA 843
47
48 74 IAsnIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThr 90
49 ::::::::::| 89
50 844 GGCGC.....ACCGGCCACTTTCGCGCTACATCTACGAAGTCCGCGC 887
51
52 91 AlaProAsnMetPheAsnValAspAsp..ValLeuGlyValIlySerPro 106
53 ::::::::::| 105
54 888 GACACAAATTTCACGGCGCGCGCGACGCTGCATTCGATTA..... 928
55
56 107 HisProTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerG 123
57
58 928 ..... 928
59
60 123 nIleTyrGlyIleTyrArgValAsnPheGlyValIleAspGluArgLeu 140
61 ..... 139
62 929 .....CGTGGAC 935
63
64 140 IsArgAsnAlaGlyIleTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
65 ::::::::::| 155
66 936 ACTTATGCGCACATGCCGCGCGTATCCTCCGCGCGCGCTGGCCACCTA 985
67
68 157 ProAlaGlnAspGlyTyrArgValAlaGlyPheProProAspHisGlnAl 173
69 ::::::::::| 172
70 986 CCGAGC.....CGAATCTCGGCACACCGCGG...CATTCGCG 102
71
72 173 aTPArgGluGlnProTyrIleHisAlaProGlnIleCysGlyAsn 190

```



```

1021 CCGAAAAACAT.....CCCCAGGCTAACCGCGGTC 1049
190 erserAgtThr.lleThrglAspThrCysasnGlutThrGlnasn 205
    |||||.....|||
1050 TATACACAGCGCATCCGCGGAGACCAACGACGAGATATTCAC 1097
seq_name: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-514-000-3096
seq_documentation_block:
; Sequence 3096, Application US/09514000
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)B
; CURRENT APPLICATION NUMBER: US/09/514,000
; CURRENT FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 15034
; SEQ ID NO 3096
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-514-000-3096

```

alignment_scores:

```

Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

```
US-09-528-682-1 x US-09-514-000-3096 ..
```

```
Align seg 1/1 to: US-09-514-000-3096 from: 1 to: 1671
```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleYr 17
  |||||.....|||
711 AACGGCAAAACCTTCATCCGACGCTTCCAGCGCCAGCTTCCAG 760
17 sArgSerGlyLeu..MetProArgGlyHisAsnGluTyrPheAsp 33
  : |||||.....|
761 CTTTCAGGCGGCTTTCAGCGCGCGGCGGCAAT.....A 798
33 rg.glyThrglMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrgl 49
  |||||.....||
799 GCGCGTGTATGCGCGCATATCGGCATC.....CGGACGCTTCA 836
49 nThrglYpHeValArgTyrAspAspGlyTyrValSerThSerThSerLeuSerL 66
  | ||| :
837 ATTCGGGTGTTTTTCAG.....CTCCGCG 859
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyr..... 80
  : |||||.....|
860 AATTCACACCTGACCTGCGGCGGCGAGTTCTTCTGCGCGGTTTCATGAC 909
81 .....SerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMe 94
  : |||||.....|
910 GAGATCTTGATGCGCGCTGTATTTGTGTGTCCGCGCGCGCGACGT 959
94 tPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG 111
  : |||||.....|
960 TATTAACGCGCATGACGCGCGGTATAGACACAGGTCCTCTTCCGCG 1009
111 lngIuValSerAlaLeuGlyGly.....IleProTyr 121
  |||||.....|
1010 AGGAGTGGCAATCTTCTTGGCATGATGCGCGCTACGCGCTCCCGCC 1059
122 SerGlnIleTyrGly.....TrrTyrArgVal 130
  : |||||.....|
1060 GCGACGTAACGCGCAACGACATCTCACGCTTGTCTTCTTTCAGAT 1109
130 lAsnPhcGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147

```

```

1110 GCAGGCGCATATCTGTGAC.....CTGAATGGCGCGCGTCCGCTCG 1153
147 sPArgTyrTyrArgAsnLeuAsnIleAlaProAlaGlnAspGlyTyrArg 163
  : |||||.....|
1154 GCACCCGCTCACCGGATGATCTTCGCGCGGACGAGACAGAAATCCGG 1203
164 LeuAlaGlyPheProProAspHisGln.....AlaThrArgGluG 177
  : ||| ||| :
1204 ATGCACGGAAGACCATGTGACGCGAGATTTCCGATAGGACGCGGATCGG 1253
177 uProTPrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThr 194
  : |||||.....|
1254 CGACTTCGTC.....TCGCGACGACCCCGGAAATGA 1285
194 lerThrglAspThrCysasnGluGluThrglAsnLeuSerThrIleTyr 210
  || |||.....|||
1286 TCGGCGGTCACTGCGCATATCGAGTTACCGGACGTCTGCAGTGCAGT 1335
211 LeuArg.....GluTyrGlnSerLysValLysArg 220
  : |||||.....|
1336 TTCACGCTTGCCAGTTCAGAGAGATATCCGCGT 1371

```

```
seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-739-449-5019
```

```
seq_documentation_block:
; Sequence 5019, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 5019
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-5019

```

```
alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

```
US-09-528-682-1 x US-09-739-449-5019/rev ..
```

```
Align seg 1/1 to reverse of: US-09-739-449-5019 from: 1 to: 1671
```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleYr 17
  |||||.....|||
961 AACGGCAAAACCTTCATCCGACGCTTCCAGCGCGGCTTTCACAG 912
17 sArgSerGlyLeu..MetProArgGlyHisAsnGluTyrPheAsp 33
  : |||||.....|
911 CTTTCAGGCGGCTTTCAGCGCGCGGCGGCGCAAT.....A 874
33 rg.glyThrglMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrgl 49
  |||||.....|
873 GCGCGTGTATGCGCGCATATCGGCATC.....CGGACGCTTCA 836
49 nThrglYpHeValArgTyrAspAspGlyTyrValSerThSerThSerLeuSerL 66
  | ||| :
835 ATTCGGGTGTTTTTCAG.....CTCCGCG 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyr..... 80
  : |||||.....|
812 AATTCACCTGACCTGCGGCGGCGAGTTCTTTCGCGCGGTTTCATGAC 763

```



```

81 .....SerThrTYRTrIleTYrValIleAlaThrAlaProAsnMe 94
762 GAGATCTTGATGCGCCCTTGATTTGTGTGCGGGCGCGCTGCAGGT 713
94 tPheAsnValAsnAspValLeuGluValTYrSerProHisProtyrGluG 111
712 TATTAACGGCGCATGACGGCGGTGATAGACACAGAGTCTCTCCGGC 663
111 lngIuValSerAlaLeuGlyGly.....IleProtyr 121
662 AGGAAGTCGGGAATCTTGTGGCATGATGCGCGCTACGGCCCTCCGCC 613
122 SerGlnIleTyrgly.....TrpTYrArgVal 130
612 GCCGAGTAACGGCAAAACCGATTCACCCCTGCTCTTTTTCAGAT 563
130 lAsnRheGluValIleAspGluArgLeuHisArgAsnArgGluTYrArg 147
562 GCAGGCGCATATCGTGAC.....CTGAATGGCGCGCGCGCTCCGCTCG 519
147 sPArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArg 163
518 GCACCCGCTACCGCGATCTTGAACTGCGCGCGAGAAACGAAATTCGG 469
164 LeuAlaGlyPheProProAspHisGln.....AlaTrpArgGluG 177
468 ATGCACGGAAAGACCATTCAGCGAGATTCGCCATAGGAGCGCGATCGG 419
177 uProTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgTrpI 194
418 CGACTTCGTC.....TCGGGAGCACACCGCGCAAAAGA 387
194 lethGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTYr 210
386 TCGGCGGTACAGTTCGAAATGCACTTACCCGACGTCGTCAGTGCAT 337
211 LeuArg.....GluTYrGlnSerLysValLysArg 220
336 TTCACGCTTGCCAGTTTCAGAGAGATATCCGGCGCT 301

```

seq_name: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-803-110-5019

```

seq_documentation_block:
; Sequence 5019, Application US/09803110
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)D
; CURRENT APPLICATION NUMBER: US/09/803,110
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: US 60/168,139
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 5019
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-803-110-5019

```

alignment_scores: Quality: 96.00 Length: 246

Percent Similarity: 0.768 Gaps: 12

Percent Identity: 50.813 Percent Identity: 22.358

alignment_block: US-09-528-682-1 x US-09-803-110-5019/rev ..

Align seg 1/1 to reverse of: US-09-803-110-5019 from: 1 to: 1671

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
961 AACGGGCAAAACCTTCATCCGACGCTTCCAGCCGCGCTTTCACAG 912
17 sArgSerGlyLeu...MetProArgGlyHisAsnGluTYrPheAsp 33
911 CTTTCAGGCGGCTCTTCAGCGCGCGGCGGAAT.....A 874
33 rg.clyThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgIleThrG 49
873 GCGCGTATGCGCGCAATATGCGCATC.....CGGCAAGCTTCA 836
49 nThrGlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSer 66
835 ATTCGGTGTTTTTCAG.....CTCCCGC 813
66 euArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYr..... 80
812 AATTCAACTGCGACCTGGCGGCGCAGTTCTTCGCGCGGCTCATGTAC 763
81 .....SerThrTYRTrIleTYrValIleAlaThrAlaProAsnMe 94
762 GAGATCTTGATGCGCGCTTGATTTGTGTGCGGGCGCGCTGCAGGT 713
94 tPheAsnValAsnAspValLeuGluValTYrSerProHisProtyrGluG 111
712 TATTAACGGCGATGACGGCGGTGATAGACAGAGTCTCTCCGGC 663
111 lngIuValSerAlaLeuGlyGly.....IleProtyr 121
662 AGGAAGTCGGGAATCTTGTGGCATGATGCGCGCTACGGCCCTCCGCC 613
122 SerGlnIleTyrgly.....TrpTYrArgVal 130
612 GCCGAGTAACGGCAAAACCGATTCACCCCTGCTCTTTTTCAGAT 563
130 lAsnRheGluValIleAspGluArgLeuHisArgAsnArgGluTYrArg 147
562 GCAGGCGCATATCGTGAC.....CTGAATGGCGCGCGCTCCGCTCG 519
147 sPArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArg 163
518 GCACCCGCTACCGCGATCTTGAACTGCGCGCGAGAAACGAAATTCGG 469
164 LeuAlaGlyPheProProAspHisGln.....AlaTrpArgGluG 177
468 ATGCACGGAAAGACCATTCAGCGAGATTCGCCATAGGAGCGCGATCGG 419
177 uProTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgTrpI 194
418 CGACTTCGTC.....TCGGGAGCACACCGCGCAAAAGA 387
194 lethGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTYr 210
386 TCGGCGGTACAGTTCGAAATGCACTTACCCGACGTCGTCAGTGCAT 337
211 LeuArg.....GluTYrGlnSerLysValLysArg 220
336 TTCACGCTTGCCAGTTTCAGAGAGATATCCGGCGCT 301

```

seq_name: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-60-168-139-1915

```

seq_documentation_block:
; Sequence 1915, Application US/60168139
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(15490)A
; CURRENT APPLICATION NUMBER: US/60/168,139

```



```

; CURRENT FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 3432
; SEQ ID NO 1915
; LENGTH: 4199
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-60-168-139-1915

```

```

alignment_scores:
  Quality: 96.00      Length: 246
  Ratio: 0.768        Gaps: 12
  Percent Similarity: 50.813  Percent Identity: 22.358

```

alignment_block:

US-09-528-682-1 x US-60-168-139-1915/rev ..

Align seg 1/1 to reverse of: US-60-168-139-1915 from: 1 to: 4199

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
   |||||
1478 AACGGCAAAACCTTCATCCGACCGCTTCAGCGCGCCAGCTTCCAG 1429
   :
17 SarSerGlyLeu...MetProArgGlyHisAsnGluTyrPheAsp 33
   :
1428 CTTTCAGGCGGCTTCAGCGCGCGCGGCAAT.....A 1391
   :
33 rg.glyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrG 49
   |||||
1390 GCGGTGATGGCGGCAATATCGGCATC.....CGGCACTTCA 1353
   :
49 nThrGlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerL 66
   |||||
1352 ATTCGCTGTTTTTCAG.....CTCCGCG 1330
   :
66 euArgSerAlaHisLeuAlaGlyInsrlIleLeuSerGlyTyr..... 80
   :
1329 AATTCAACCTCGACCTGGCGGCGCATCTTCGCGCGGTTTCATGTAC 1280
   :
81 .....SerThrTyrTyrIleTyrValIleAlaThrAlaProAsn 94
   :
1279 GAGATCTGTATGGCGCGCTGTATTTGTGTGCGCGCGCGCTGCAGAT 1230
   :
94 rPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGln 111
   :
1229 TATAAACCGCATGACGCGCGCTGTATGACAGACAGAGCTCTCCGCG 1180
   :
111 lngIuValSerAlaLeuGly.....IleProTyr 121
   :
1179 AGGAGTGGCGAATCTTGTGCGCATTCAGCGCTACGCGCTGCCGCC 1130
   :
122 SerGlnIleTyrGly.....TyrTyrArgVal 130
   :
1129 GCGGACGTAACGCGAAGCCGATCTGCAGCGCGGAGAACGAAATGCCG 986
   :
130 lAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
   :
1079 GCAGGCGCATGTCTGTGAC.....CTGATGGCGCGCGCGTCCGCGT 1036
   :
147 sPArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg 163
   :
1035 GCACCGCTGACCGCATCTGTACACTGCGCGGAGAACGAAATGCCG 986
   :
164 leuAlaGlyPheProProAspHisGln.....AlaTrpArgGluG 177
   :
985 ATGCAAGCAAGACCATTTGACGAGATTTCCGCAATAGGACCGCGATGCG 936
   :
177 uProTyrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThrI 194
   :
935 CGACTTCGTC.....TGGCGGACGACCGCGGCAATGA 904
   :
194 lThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
   :

```

```

903 TCGGCGGTACAGTTGCGAATGACAGTACCGGACGCTTCAGTGGCTGCAT 854
211 leuArg.....GluTyrGlnSerLysValLysArg 220
      :
853 TTCACGCTTGCGAGTTCAGAGAGGATATCCGGCGT 818

```

seq_name: /cgn2.6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-45559

seq_documentation_block:

; Sequence 45559, Application US/09620392

; GENERAL INFORMATION:

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong

; APPLICANT: McIninch, James

; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof

; FILE REFERENCE: 38-21(51237)E

; CURRENT APPLICATION NUMBER: US/09/620,392

; CURRENT FILING DATE: 2000-07-19

; NUMBER OF SEQ ID NOS: 69652

; SEQ ID NO 45559

; LENGTH: 7412

; TYPE: DNA

; ORGANISM: Oryza sativa

US-09-620-392-45559

```

alignment_scores:
  Quality: 96.00      Length: 195
  Ratio: 1.103        Gaps: 9
  Percent Similarity: 44.615  Percent Identity: 24.103

```

alignment_block:

US-09-528-682-1 x US-09-620-392-45559 ..

Align seg 1/1 to: US-09-620-392-45559 from: 1 to: 7412

```

2 GlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeuSar 18
   |||||
245 GGTGTCAGGCTACGAGAACGATTCACAG.....AAGAG 279
   :
18 gSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPhe..... 31
   :
280 AAGAAGAGGTGGTGGTGGCGCGCGCGCATGATACCTTGTAGCTGGCG 329
   :
32 .....AspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArg 46
   :
330 ACCCAGATGTGCGCGTCAAGTACAGACCTGGCGGAGCGGACATGCTCG 379
   :
47 glyThrGlnThrGlyPheValArgTyrAsp.....AspGlyTyrValSe 61
   :
380 GGCACCGTGTATGGCGCTGGCGGAGCGGACGATCTTACAGGTGAATCGC 429
   :
61 rThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
   :
430 CTTGCGCGTCCCGCTGCGCGCGCGGCGGACGAGCTGGCGAG..... 469
   :
78 ergIuTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMet 94
   :
469 ..... 469
95 PheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGlnG 111
   :
470 .....CTTCACATGCGGCGTTCGCA 489
   :
111 nGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrP 128
   :
490 GCGGCTCACGAGGTGCGCGCGCTC.....GAGGTGCTGCGG... 526
   :
128 yArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlu 144
   :
527 .....GATCTCGAGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 533

```



```

111 ngIuVaISeraLaLeuGLyIleProTyrSerGlnIleTyrGLyTrpT 128
:::      ::::|:::      ::::  ||

```

```
94 rPheasnaValAsnAprValLeuGlyValTySerProHisProGlyGlu 111
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```



```
6352 TATAACGCCGATAGCGCCGTGTTATAGACACAGAGTCTCTCCGCC 6401
111 InGUValSerAlaLeuGly.....IleProTyr 121
|||||.....|||
6402 AGGAAGTCGGGAATCTTGTGGCATCATGGCGGTACGCCCTGCCGCC 6451
122 SerGlnIleTyrGly.....TrpTyrArgVa 130
.....:|
6452 GCCGACGCTAACGCGCAAAACCGATCTCACCCCTGTCTCTTTTCAGAT 6501
130 LasnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
.....:|
6502 GCAGGCCGATATCGTGAC.....CTGAATGGGGGGCGCGCTCGCCTCG 6545
147 sParGlyTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg 163
:::|||||:|
6546 GCACCCGTCACCGGATCTTGAACTGCGGCGCGAAGAGAAATTCGG 6595
164 LeuAlaGlyPheProAspHisGln.....AlaTPrArgGluG 177
::|:|
6596 ATGCACCGAAGACCATGACGACGATTCGCGATAGGACGCGGATCGG 6645
177 uProTPrIleHisIleAlaProGlnGlyAsnSerSerArgThr 194
.....:|
6646 CGACTTCGTC.....TGCAGCAGCACCCGCAAAATGA 6677
194 leThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
|||:|:|
6678 TCGGCGGTCAAGTTGCGAATGCAATTACCGACGTCGTCAGTGCGTCAT 6727
211 LeuArg.....GluTyrGlnSerLysValLysArg 220
.....:|
6728 TTCACGCTTGCCAGTTCAGAGAGATATCCGGCGT 6763
```


[illegible]

/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-105-289-7102	-	71.00	111.93	403.03	
/cgn2_6/ptodata1/pna/PCrT_NEM_COMB.seq:PCrT-US02-US02-9288-34	-	71.00	111.18	443.24	
/cgn2_6/ptodata1/pna/PCrT_NEM_COMB.seq:PCrT-US02-US02-9288-35	-	71.00	110.52	482.56	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-10-09-667-170A-422	-	71.00	100.17	1.8e+03	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-105-299-11480	-	71.00	88.17	8.5e+04	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-216283	+	70.50	128.10	50.6	128.10
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-540-209B-4950	+	70.50	121.49	118.2	121.49
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-44352	+	70.50	120.91	127.3	120.91
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-919-002-9441	+	70.50	120.84	128.4	120.84
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-40452	+	70.50	118.20	180.1	118.20
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-40734	+	70.50	117.96	185.6	117.96
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-28551	+	70.50	116.70	218.6	116.70
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-104-047-11307	+	70.50	114.89	275.5	114.89
/cgn2_6/ptodata1/pna/PCrT_NEM_COMB.seq:PCrT-US02-1314-568	-	70.50	110.72	470.53	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-128-714-5568	-	70.50	110.72	470.53	
/cgn2_6/ptodata1/pna/PCrT_NEM_COMB.seq:PCrT-US02-1314-5568	-	70.50	110.68	472.5	110.68
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-128-714-5568	-	70.50	110.68	472.5	110.68
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-088-697-1	+	70.50	108.08	659.91	
/cgn2_6/ptodata1/pna/US08_NEM_COMB.seq:US-08-793-273C-1	+	70.50	104.36	1.1e+03	
/cgn2_6/ptodata1/pna/US08_NEM_COMB.seq:US-09-919-002-9441	+	70.50	104.22	1.1e+03	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-100-053-375B-1354	+	70.50	103.99	1.1e+03	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-100-912-913	+	70.50	103.99	1.1e+03	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-935-625-11845	+	70.50	101.04	1.6e+03	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-935-625-26236	+	70.50	96.10	3.1e+03	
/cgn2_6/ptodata1/pna/US08_NEM_COMB.seq:US-09-503-138B-1	+	70.50	96.10	3.1e+03	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-105-407-1	+	70.50	96.10	3.1e+03	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-76212	-	70.00	128.07	50.6	128.07
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-242617	-	70.00	127.14	57.2	127.14
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-242617	-	70.00	123.62	89.8	123.62
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-395958	+	70.00	123.62	89.8	123.62
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-395951	+	70.00	123.39	92.6	123.39
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-36747	+	70.00	123.25	94.26	
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-10-152-886-52	+	70.00	122.81	99.78	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-339206	+	70.00	121.92	111.1	121.92
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-254926	+	70.00	121.92	111.1	121.92
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-35130	+	70.00	117.27	203.1	117.27
/cgn2_6/ptodata1/pna/US06_NEM_COMB.seq:US-60-360-033-46935	-	70.00	116.83	214.8	116.83
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-104-047-1125	+	70.00	114.19	301.55	114.19
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-982-667-2	+	70.00	114.02	308.06	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-982-667-2	+	70.00	114.02	308.06	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-033-297-2	+	70.00	114.02	308.06	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-540-209B-4067	+	70.00	112.60	369.5	112.60
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-105-299-13682	-	70.00	99.16	2.1e+03	
/cgn2_6/ptodata1/pna/US09_NEM_COMB.seq:US-09-939-964A-1	-	70.00	60.41	2.8e+03	
/cgn2_6/ptodata1/pna/US10_NEM_COMB.seq:US-10-027-633-277332	+	69.50	126.89	59.1	126

[illegible]

[illegible]

[illegible][illegible]

/cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-6307 + 64.50 116.54 222.93
 /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-149787 - 64.50 114.79 279.14
 /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-33329 + 64.50 112.99 351.63
 /cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-33742 - 64.50 112.37 380.89
 /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-704-302A-280 + 64.50 112.35 381.92
 /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3716 + 64.50 112.24 387.11
 /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-11904 - 64.50 112.19 389.71
 /cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41713 + 64.50 112.08 394.93
 /cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-43416 + 64.50 112.05 396.49
 /cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37252 + 64.50 110.57 479.75
 /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-122562 + 64.50 109.23 569.70

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-1

seq_documentation_block:

; Sequence 1, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McChee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshitumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; REFERENCE/DOCKET NUMBER: MCG-01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

; US-09-809-033A-1

alignment_scores:
 Quality: 1053.00 Length: 240
 Ratio: 4.701 Gaps: 0
 Percent Similarity: 93.33 Percent Identity: 80.833

alignment_block:

US-09-528-682-1 x US-09-809-033A-1 ..

Align seq 1/1 to: US-09-809-033A-1 from: 1 to: 2022

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeuLeu 17

|||||
 572 AAGCATCATTAAGTTATATCGGACGATTTCTAGACCTTCATCAATAATAA 621
 17 sArgSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||
 622 GCGATCAGGTGGTCTTATGCCACAGACACAGATGACTTTGACCGAG 671
 34 lYhrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 672 GTACTCAATAGATATCAACCTTATGATCATGCACAGAGAACTCAGACG 721
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 722 GGATTTGTTAGGCACGATGATGATATTTCTCACCCTCAATAGTTGAG 771
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 772 AAGTCCCACTTATGTTGGTCAACTATATTGTGTGCTCATTTACTTATT 821
 84 yrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 822 ATATATATATGTTATAGCCACGACCCACACATGTTTAACGTTATATGATGA 870
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
 871 TTAGGGGCATACAGCTCCTCATCCAGATGACAGAAAGATTCTCTTAGG 920
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAspGlyVal 134
 921 TGGGATTCATCTCCCAATATATGATGATGATTCGATTCATTTTGGG 970
 134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 971 TGCTGTATGAACAATTATCATCGTAATAGGGGCTACAGAGATATATATTC 1020
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 1021 AGTAACCTTATATTTGCTCCAGACGACGAGATGATGATGATGGCAGATT 1070
 167 eProProAspHisGlnAlaIleThrArgGluGluProTyrPheHisAla 184
 1071 CCTCCGACCATAGACCTTGGAGGACGACCGCTGATTCATCATTCAC 1120
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 1121 CGCGGGTGTGGGAATGCTCAAGATCATCGATGATATCTTCGAT 1170
 201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 1171 GAAAAACCCAAAGCTAGGTGTAATAATCTTGACGAAATCAACAACTCA 1220
 217 sValIlysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 1221 AGTTAAAGACAAATATTTTTCAGGCTATCATCTGATATTGATACACAT 1270
 234 snArgIleArgAspGluLeu 240
 1271 ATGAATTAAGATGATTA 1290

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-2

seq_documentation_block:

; Sequence 2, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McChee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshitumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4


```

33  rg.glyThrGlnmetAsnIleasnLeuylxrsphIsAlaAargylThrGl 49
34  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
873  GCGCGTGATGCGCGCAATATCGGCATC.....CGGCAGCTTCA 836
49  nThrGlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerL 66
50  | ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
835  ATTCGGTGTGTTTTCAG.....CTCGCG 813
66  euArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyr..... 80
67  :::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
812  AATTCACCTCGACCTGCGCGCGAGTTCTTGGCGCGGTTTCATGATGAC 763
81  .....SerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMe 94
762  GAGAAATCTTGATCGCGCCTGTGATTGTGTGTGGGCGCGCGTGCAGGT 713
94  rPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGlnG 111
95  ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
712  TATAACGGCGCATGACGCGCGGTATGACAGAGGTCTCTTCGCGC 663
111  IngIValSerAlaLeuGlyLys.....IleProTyr 121
112  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
662  AGCAAGTGGCGCAATCTCTTGGCGATCGACGCGCTGCGCGCGC 613
122  SerGlnIleTyrGly.....TyrIleArgVal 130
612  GCGGACGTAACGCGCAAAACGATCTCACCTTGTCGTTCTTTTCACAT 563
130  IAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
131  :::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
562  GCAGCGCCGATTCGTGAC.....CTGAATGCGCGCGCGTCCGCGTGC 519
147  sPArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg 163
518  GCACCCGTCACCGCGATCTTAACCTTGGCGCGCAGAAACGAAATTCGG 469
164  LeuAlaGlyPheProProAspHisGln.....AlaTyrArgGluG 177
165  ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
468  ATGCACGGAAGACCATTCGACGAGATTTCCGCATAGGACGCGGATCCG 419
177  uProTyrIleHisIleAlaProGlnGlyCysGlyAsnSerSerArgThrI 194
418  CGACTTCGTC.....TGGCGGACGACCCGCAAAATGA 387
194  IeThrGlyAspPheCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
211  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
386  TCGCGGCTCAGCTTCGAAATGACATTACCGACGCTGCAAGTGCATGCG 337
211  LeuArg.....GluTyrGlnSerLysValIleAspArg 220
336  TTCACGCTTGCCAGTTTCAGAGAGATATCGGCGGT 301
seq_name: /cgn2_6/plotdata/1/pna/US60_NEW_COMB.seq:US-60-360-039-38194
seq_documentation_block:
: Sequence 38194, Application US/60360039
: GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
NUMBER OF SEQ ID NOS: 2002-02-21
SEQ ID NO 38194
LENGTH: 1653
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens

```

[illegible]


```

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-38598
seq_documentation_block:
: Sequence 38598, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 38598
: LENGTH: 1653
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-360-039-38598

```

```

alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

US-09-528-682-1 x US-60-360-039-38598/rev ..

Align seg 1/1 to reverse of: US-60-360-039-38598 from: 1 to: 1653

```

1 ASnglYAspArgLeuTYrArgAlaAspSerArgProProAspLulely 17
  |||||:|||||:|||||
961 AACGGCAAAACCTTATCCGACGCTCCAGCGCGCGCTTCCAG 912
17 sArgSerGlyLeu...MetProArgLynHisAsnGluTYrPheAsp 33
  :|||||:|||||:|||||
911 CTTTCAGGGGGCTTTTCAGCGCGCGCGCGGAAT.....A 874
33 rg.GlYThrcImetAsnIleAsnLeuTYrAspHisAlaArgLylThrgl 49
  |||||:|||||:|||||
873 GGGGTGATGGCGCGCAATATCGGCATC.....CGCAGCTTCA 836
49 nThrgLyrPheValArgTYrAspAspLylTYrValSerThrSerLeuSerl 66
  |||||:|||||:|||||
835 ATTCGGGTGTTTTCAG.....CTCCGGC 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
812 AATTCAACCTCGACCTGGCGGGCGAGTTCTTGCAGCGCTTCATGATAC 763
81 .....SerThrTYrLylTYrValIleAlaThralaProAsnme 94
762 GAAATCTTGATCGCGCTTGATTTGTGTGCGCGCGCGCTTCAGGT 713
94 tPheAsnValAsnAspValLeuGlyValTYrSerProHisProTYrGlu 111
  :|||||:|||||:|||||
712 TATAAACGGCGATAGCGCGGTGTATAGACAGCAGGTCTCTCCGCG 663
111 lnglValSerAlaLeuGlyLyl.....IleProTYr 121
  |||||:|||||:|||||
662 AGGAATCTCGGAATCTTTCGGATCATCGCGGTACGCGCGCTCCGCG 613
122 SerGlnIleTYrGly.....TrpTYrArgVal 130
  :|||||:|||||:|||||
612 GCGGAGGTAAACGGCAAAACCGATTCACCTTGCTTTTTCGAT 563
130 lAsnPhgGlyValIleAspGluArgLeuHisArgAsnArgLylTYrArg 147
  :|||||:|||||:|||||
562 GCAGGGCGGATATCGTGAC.....CTGAATGCGGGCGGTGCGGCTCG 519

```

```

147 sPArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspLylTYrArg 163
  :|||||:|||||:|||||
518 GCACCCGTCACCCGCGATCTTGAATTCGGCGGACAGACAGATTCGG 469
164 LeuAlaGlyPheProProAspHisGln.....AlaTrpArgGlu 177
  :|||||:|||||:|||||
468 ATCCAGCGAAGACCATTTGACGAGATTTCCGATAGGACGCGGATTCG 419
177 uProTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgThr 194
  :|||||:|||||:|||||
418 CGACTTCGTC.....TCGGCAGCAGACCGCGGAATGA 387
194 lPheGlyAspThrcysAsnGluGluThrcGlnAsnLeuSerThrlTYr 210
  |||||:|||||:|||||
386 TCGCGGTTCAGCTTCGCAATGCAGTTACCGAGCTCTCGATCGGTGAT 337
211 LeuArg.....GluTYrGlnSerLysValysArg 220
  :|||||:|||||:|||||
336 TTCACGCTTGCAGATTTCAGAGAGATATCCGGCGT 301

```

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37962

```

seq_documentation_block:
: Sequence 37962, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 37962
: LENGTH: 1657
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-360-039-37962

```

```

alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

US-09-528-682-1 x US-60-360-039-37962/rev ..

Align seg 1/1 to reverse of: US-60-360-039-37962 from: 1 to: 1657

```

1 ASnglYAspArgLeuTYrArgAlaAspSerArgProProAspLulely 17
  |||||:|||||:|||||
961 AACGGCAAAACCTTATCCGACGCTTCAGCGCGCGCGCTTCCAG 912
17 sArgSerGlyLeu...MetProArgLynHisAsnGluTYrPheAsp 33
  :|||||:|||||:|||||
911 CTTTCAGGGGGCTTTTCAGCGCGCGCGCGGAAT.....A 874
33 rg.GlYThrcImetAsnIleAsnLeuTYrAspHisAlaArgLylThrgl 49
  |||||:|||||:|||||
873 GGGGTGATGGCGCGCAATATCGGCATC.....CGCAGCTTCA 836
49 nThrgLyrPheValArgTYrAspAspLylTYrValSerThrSerLeuSerl 66
  |||||:|||||:|||||
835 ATTCGGGTGTTTTCAG.....CTCCGGC 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
812 AATTCAACCTCGACCTGGCGGGCGAGTTCTTGCAGCGCTTCATGATAC 763

```



```

|||||
1747 GGTGTGATCGACGATACGTGAACAGTTTGATGACCAACCAATGAC 1796
141 ..... 141
1797 AAGTACCAAGGTGCAACCCAGATTATGCACACGTTCTGTATTGGAC 1846
142 ....AsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
1447 GCGTGAAGTATGATATATATCCCGTATTTATTTGAGGCAATTC... 1893
157 ProAlaGluAspGly.....TyrArgLeuAlaG1 166
1894 ...CGTTATGACGGCTCCCTCACGTTTCCACAAAGATCATCGTTGGGATT 1940
166 yPheProProAspHisGlnAlaIleTyrArg.....GluGluProThrIleH 181
1941 CTTCCCTTCCTTATCAGGTGCATGGCGCATCTCAGAAAGAACTTCATGG 1990
181 IsHsAlaProGlnGlyCysGlyAsn...SerSerArgThrIleThrGly 196
1991 AAATTAACCCGTAAGCTGGCTGGATTAACCTGAAGAAAGTCGTCATCTGGGGT 2040
197 AspHrcyAsnGluGluThrGlnAsnLeuSerThrIleTyrLeu 211
2041 AAATTAAGTAACTCTGAATTTGTAACTACGAAATACATGTGGGTA 2085
seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:us-60-360-039-46582

```

```

seq_documentation_block:
; Sequence 46582, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46582
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Aeropyrum pernix
US-60-360-039-46582

```

```

alignment_scores:
Quality: 86.00 Length: 187
Ratio: 0.956 Gaps: 10
Percent Similarity: 48.128 Percent Identity: 21.925

```

alignment_block:

US-09-528-682-1 x US-60-360-039-46582 ..

Align seg 1/1 to: US-60-360-039-46582 from: 1 to: 1272

```

11 ArgProProAspGluIleLysArgSerGlyGlyLeuMetProArg..... 25
19 CAGCCACGCGACCTCTTGAAGAGCTCGCGCAGCTCAGAGAGAGCTGT 68
26 .....GlyHisAsnGluTyrPhe.....A 32
69 AGACCTTCTGGGCTAGCATGACTATGTTATGAAGTTCTGCACAAACCCGG 118
32 spArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThr 48
119 ACAGAGTTCTACAGTAAGGCTCACT..... 144
49 GlnThrGlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSe 65

```

```

145 .....ATTAGATGATGACGATCTGCTCAAGACTTCCTCGG 182
65 rLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSer 82
183 CTGGAGGAGGACGACCATAGCGGCTAGGCCCCCTACAAAGGGGCTTA 232
82 hTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsn 98
233 GGTACAC.....CCCAACGTT...ACTATGAAT 258
99 AspValLeuGlyValTyrSerProHisProTyrGluGlnValSerAl 115
259 GAGGTATACCCCTTACATGATGATGATGATGATGATGATGATGATGAT 390
115 aleuGlyGlyIleProTyrSerGlnIleTyrGlyTyrPyrTyrArgValAsp 132
300 CCGCGCGGCTGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 349
132 heGlyValIleAspGluArg.....LeuHisArgAsnArgGlu 144
350 CTAAGATATCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 390
145 TyrArgAspArgTyrTyrArgAsnLeuAsn..... 154
391 .....AAGTATTCGAGAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 431
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsp 171
432 GGACATTCGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 469
171 IsGlnAlaTyr 174
470 TCATGTCTTGG 480

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:us-10-104-047-43

```

seq_documentation_block:
; Sequence 43, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 2338
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-43

```

```

alignment_scores:
Quality: 86.00 Length: 199
Ratio: 0.827 Gaps: 11
Percent Similarity: 52.261 Percent Identity: 23.618

```

alignment_block:

US-09-528-682-1 x US-10-104-047-43 ..

Align seg 1/1 to: US-10-104-047-43 from: 1 to: 2338

```

45 AlaArgGlyThrGlnGlyPheValArgTyrAspAspGlyTyrValSe 61
134 GCCCGAAGCTCGGAGATGACATCGGCGGAGCCGATACCCGGGGCAG 183
61 rThrSerLeuSerLeuArgSerAlaHisLeu.....AlaGly 74
184 GCGAGCCCTGGAGCAACCGCTACCCACCTCAGGCGGCTGGTGGG 233

```



```

74 Inse.....IleuSerGlyTyrSerThrTyrIleValIle 88
   ||| ..... ||| ..... |||
234 GCACCTCATGTGTGCTGGGGCAGC.....CTCATGTGCTG 274
   ||| ..... ||| ..... |||
89 AlaThrAlaProAsnMetPheAsnValAsnAspValIleuGlyValTyrSe 105
   ||| ..... ||| ..... |||
275 GCTGGGGGACGCCCATGTGTGCTGGGGACAGCCTCATGTGTCAGC 324
   ||| ..... ||| ..... |||
105 rProHISProTyrGluGluValSerAlaLeuGlyIleProTyrS 122
   : ..... : ..... :
325 TGGG.....GCAGTGTGGGGGCTT..... 346
   : ..... : ..... :
122 erGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArg 138
   ||| ..... ||| ..... |||
347 .....GGCGCTCATGCTGGGAAGTCGCATGACGAGCCCATC 385
   ||| ..... ||| ..... |||
139 LeuHISArgAsnArgGluTyrArgAspArgTyrTyrArg...AsnLeuAs 154
   ||| ..... ||| ..... |||
386 ATTCATTTGGCGAGTACTATGAGGACCGTACTATCGTAAACATGCA 435
   ||| ..... ||| ..... |||
154 nIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
   : ..... : ..... :
436 CCCTACCCCAACCAAGTGTACTACAG.....CCCATGGATG 473
   : ..... : ..... :
171 IsGlnAlaTrpArgGluGluProTyrIleHISHisAlaProGlnGlyCys 187
   : ..... : ..... :
474 AGTACAGC...AACCAACACTTGTGTGC.....GACTGC 508
   : ..... : ..... :
188 GlyAsn.....SerSerArgThrIleThrGlyAspThrCysAsnG 201
   ||| ..... ||| ..... |||
509 GTCAAATTCATCAAGCAGCAGCGTCACCAACCAACGAGGGA 538
   ||| ..... ||| ..... |||
201 uGluTrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLysV 218
   ||| ..... ||| ..... |||
559 G.....ACTTCACGAGACGACGCTTAAGATGATGAGCGCGCTG 599
   ||| ..... ||| ..... |||
218 AllysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 233
   ||| ..... ||| ..... |||
600 TGAGAGAGATGTATACCCAGTACGAGGAATCTCAGGCTAT 646
   ||| ..... ||| ..... |||
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:us-10-027-632-252780
seq_documentation_block:
; Sequence 252780, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218.006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198.676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193.483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185.218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167.363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156.358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146.002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252780
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-252780

```

```

alignment_scores:
  Quality: 85.00      Length: 221
  Ratio: 0.833      Gaps: 10
  Percent Similarity: 46.154      Percent Identity: 19.457
alignment_block:
US-09-528-682-1 x US-10-027-632-252780/rev ..
Align seg 1/1 to reverse of: US-10-027-632-252780 from: 1 to: 2298

51 GlyPheValArgTyrAspAspGlyTyrValSerThr.....SerLe 64
   ||||| ..... ||| ..... |||
934 GGTTTTACATGAATTCATCTGAGTATCTAGAGACTTACTGCTTTTCT 885
   ||||| ..... ||| ..... |||
64 uSerLeuArgSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrS 81
   ||||| ..... ||| ..... |||
884 CTCCTCTCTCTCTCTCTCTCTCTGTCAGCTTGAGCTTTGAGTACC 835
   ||| ..... ||| ..... |||
81 erThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPhe..... 95
   : ..... : ..... :
834 AG.....GCCACTGACCACATCTCTTTAGATTC 806
   : ..... : ..... :
96 .....AsnValAs 98
   : ..... : ..... :
805 TCTCCAAAGATATTGAAGAGATCGTCATCCATGCTATACAAATATC 756
   : ..... : ..... :
98 nAspValIleuGlyValTyrSerProHISProTyrGluGluValSerA 115
   :||| ..... ||| ..... |||
755 AGATCATTTATCTGTCCACCTCTCA..... 731
   : ..... : ..... :
115 lAleuGlyGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsn 131
   ||| ..... ||| ..... |||
730 .....CCCCAAACTAACTCTCAAGTTTCTGAATTTAAA 695
   : ..... : ..... :
132 PheGlyValIleAspGluArgLeuHISArgAsnArgGluTyrArgAspAr 148
   : ..... : ..... :
694 AAAAATTCATTTACCAACACATGCAT...CATAGTATCATGAGCAAA 648
   : ..... : ..... :
148 gTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuA 165
   : ..... : ..... :
647 TGAGATTAAAGATGCAATGCGTATCTCAAGAGCTCATATTCATTT. 599
   : ..... : ..... :
165 lAglyPheProProAspHisGlnAlaTrpArgGluGluProTyrIleHIS 181
   ||||| ..... ||| ..... |||
598 .....TGAGAGAACACATATCATGCAC 575
   : ..... : ..... :
182 ...HisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAs 197
   ||| ..... ||| ..... |||
574 GTCACACA.....TGCACAAATGTTCAAGACACCAACCAAGA 534
   : ..... : ..... :
197 PThr..... 198
   : ..... : ..... :
533 GACCCATCAAGGCTTCATGTCCTTTTCCAAATGATGAGCAGGAT 484
   : ..... : ..... :
199 .....CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArg 212
   ||| ..... ||| ..... |||
483 AGCTGATTTGTTAAACGAGGAGAACTTACAC...CTGTGTCGAGA 437
   : ..... : ..... :
213 GluTyrGlnSerIysValLysArgGlnIlePheSerAspTyrGlnSerG 229
   : ..... : ..... :
436 GAGAGAAAGACAGCAGTAAACAAAGAGACTATGATAGCTAGTGTACA 387
   : ..... : ..... :
229 uValAspIleTyr 233
   ||| ..... ||| ..... |||
386 GGTCAATTCGAT 374
   : ..... : ..... :
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-8511
seq_documentation_block:
; Sequence 8511, Application US/09919002
; GENERAL INFORMATION:

```



```

; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
; PRIOR FILING DATE: FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8511
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-002-8511

```

```

alignment_scores:
  Quality: 84.00      Length: 168
  Ratio: 0.977        Gaps: 10
  Percent Similarity: 51.190      Percent Identity: 23.810

```

alignment_block:
US-09-528-682-1 x US-09-919-002-8511/rev ..

Align seg 1/1 to reverse of: US-09-919-002-8511 from: 1 to: 858

```

53 ValArgTyrAspAspGlyTyrValSerThrSerLeuArgSerAl 69
|||||:|||||:|||||:|||||:|||||:
652 GTTCGGTTTCATGCGCTGATATCATCAAAACCATATATTCGTCAA.. 605
69 ahIsleuAlaGlyInSerIleuSerGlyTyr..... 80
|||||:|||||:|||||:|||||:|||||:
604 .....GGGGACAGCTCTTGTATGTTTATAGAGCCCTGGACCC 565
81 ..SerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsn 96
|||||:|||||:|||||:|||||:
564 AAGTGAATATTACAGGTACATA.....AGATTCTTCCCT 530
97 ValAsnAspValLeuGlyValTyrSerProHisProTyrGluGluVal 113
|||||:|||||:|||||:|||||:|||||:
529 GATGGCCATGTGATGATGTGACAAACCCGTGAAGAGCCTCAGTCCATTGT 480
113 lSerAlaLeu...GlyGlyIleProTyrSerGlnIleTyrGlyTyrTyr 128
|||||:|||||:|||||:|||||:|||||:
479 TTCACGTTTAAAGACTAGAGATACAGACTGATGCAATTCTCTGCTC 430
129 ArgValAsn.....PheGlyVal 134
|||||:|||||:|||||:|||||:
429 ACTATTCCTGTGTCAAGACACAGACACATCAGACCAAAAGATTGCTGT 380
134 lIleAspGluArgLeu...HisArgAsnArgGlyTyrArgAspArgTyr 150
|||||:|||||:|||||:|||||:|||||:
379 AATTAACCTAAGAAAAAGAGAAAAACCACTGACTATTAATATACAGATTT 330
150 YrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg.....Leu 164
|||||:|||||:|||||:|||||:|||||:
329 TTCGTGCGTCCCTGTACAGAGACAGATCAGAGTTTTCATGTGGGCTA 280
165 AlaAlaGlyPheProAspHisGlnAlaTyrArgGluLuprotRPIleH 181
|||||:|||||:|||||:|||||:|||||:
279 CAGCTATGTCCAGTGGTACACAGAGGTTCACAAACTCATCTGGATACA 230
181 SHSAlaProGlnGlyScgIlyAsnSerSerArgTThrIleThrgIlyAsp 198
|||||:|||||:|||||:|||||:|||||:
229 TCATTCT.....TGTCACATTACTTACAATCA...ACTGGTGA 192
198 hr 198
||
191 CT 190

```

seq_name: /cgn2.6/plodata/1/pna/US09_NEW.COMB.seq:US-09-919-002-6157

```

seq_documentation_block:
; Sequence 6157, Application US/09919002
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
; PRIOR FILING DATE: FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6157
; LENGTH: 3617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-002-6157

```

```

alignment_scores:
  Quality: 83.00      Length: 228
  Ratio: 0.761        Gaps: 16
  Percent Similarity: 47.807      Percent Identity: 22.368

```

alignment_block:
US-09-528-682-1 x US-09-919-002-6157 ..

Align seg 1/1 to: US-09-919-002-6157 from: 1 to: 3617

```

40 AsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAs 56
|||||:|||||:|||||:|||||:|||||:
2132 CACCTTACTGATACATGAGCATGTACTAGAGAAAAACCTGCAAAATATGA 2181
56 pAsp.....GlyTyrValSer...ThrSerLeuSerLeuArgSerA 69
|||||:|||||:|||||:|||||:|||||:
2182 AGATGCGACAGTGTCTTATTCCTTGTGCTCAGACGCTTCTAATCAACAG 2231
69 lAhIsleuAlaGlyInSerIleuSerGlyTyrSerThrTyrTyrIle 85
|||||:|||||:|||||:|||||:|||||:
2232 TGATTCTTGTGAGAGAGAG..CTCTCCAAATGTGAACATGTGTC... 2275
86 TyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeuG 102
|||||:|||||:|||||:|||||:|||||:
2276 .....AAAGTTTAAAC..... 2287
102 yAlaTyrSerProHisProTyrGluGluGluValSerAlaLeuGlyTyr 119
|||||:|||||:|||||:|||||:|||||:
2288 ....CACAGCCAAATCTCTTCCAAACACAGAGAAATGAGATGGAGGGA 2333
119 lProTyr.....SerGlnIleTyrGlyTyr..... 128
|||||:|||||:|||||:|||||:|||||:
2334 AACCTTCAATGTGAGAGATGTGACAGCATCTTCAAGTGTCTCAGAC 2383
129 .....ArgValAsnPheGlyValIleAspGluArgLeuH 140
|||||:|||||:|||||:|||||:|||||:
2384 CTTACTTAACATTAAGAAATTCACACTGTT.....GAGAACCATA 2424
140 sArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
|||||:|||||:|||||:|||||:|||||:
2425 CAATGTGACGAATGTGGGAAGCCTATATACACAGTCTCACACCTCAGTG 2474
157 .....ProAlaGluAspGlyTyrArg..... 163
|||||:|||||:|||||:|||||:|||||:
2475 AACACAGAGAGATTTCACACCGGAGAGAAACCTTACCATGTGAAGATGT 2524
164 .....LeuAlaGlyPheProAspHis 171

```



```

2525 GCGAAGCTCTTCAGAACTTGTCTCAGCCCTTCTACACATAAGAACTCA 2574
171 sglalatrpargluInuProtrilEhshslaprogInlycsc 188
2575 TTCTT.....GAAGAAAAACC.....TACACCTGTGAAGAAATGTG 2609
188 lysn.....SerSerArgThrIle 194
2610 GCAACATCTTTAAGCAGTATACAGACTCCTACATAAGAAAAACCAT 2659
195 ThrGly.....AspThrCysasnGlu.....ThrG 204
2660 ACTGAGAGAGAACCCCTCAAAATGTGACGATGTGAAAAAATTACCA 2709
204 nasLeuSerThrIle.TyrLeuArgGluTyr 214
2710 GTCCTCCACCTTATGTATACATAAGAAATTC 2741

seq_name: /cgn2_6/ptodata/1/pna/us60_NEW_COMB.seq:us-60-360-039-32681

seq_documentation_block:
; Sequence 32681, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 32681
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Chloroflexus aurantiacus
US-60-360-039-32681

alignment_scores:
Quality: 82.50 Length: 219
Ratio: 0.764 Gaps: 12
Percent Similarity: 49.315 Percent Identity: 26.027

Alignment_block:
US-09-528-682-1 x US-60-360-039-32681 ..
Align seg 1/1 to: US-60-360-039-32681 from: 1 to: 1107

3 AspArgLeuTyrArgAlaAspSerArgProProAspGluIleLysArgSe 19
141 GACCGCATTCGCGATGACGACCGCAGCCACCTCAGCAGACCGCGGATC 190
19 rGlyGlyLeuMetProArgGlyHisn.....G 29
191 G.....CGTTCACACGACCACTGTCCATCAGTGAAACAG 225
29 LuTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAla 45
226 AGCAAAATGCTTCGCGAGAACTGAAACCGACTGTACAGCTGGATGAGAA 275
46 Arg.....GlyThrGln..ThrGlyPheValA 54
276 CGGAATCGAACCGCGTATGTCTCTGCGCCCTGCGTGGCGGCGGCG 325
54 rGlyTyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHis 70
326 GTGACCTGCGAGCGGTTCGCGCAGCAGCAATC..... 357
71 LeuAlaGlyInsIleLeuSerGlyTyrSerThrTyrTyrIleTyrVal 87

```

```

358 TTGGCGCGGTGTGGCCCTCTCCATACACCGACGACGCTCCGGGATGCT 407
87 lIleAlaThrAlaProAsnMetPheAsnValAsnAspValIleGly.... 102
408 TGATGCTGCGGTAGCTGTAAACCGGCATCAGCACCCGCTGGGCAAA 457
103 .....ValTyrSerProHisProTyrGluInGluValSer 114
458 ACCTGATCGGAGCCTTTCATCAACACGCGTGTGTGGCCGACACCAAC 507
115 AlaleuGlyGlyIleProTyrSerGlnIle...TyrGlyTyrTyrArgVal 130
508 CTGTGGCGACACTACCGCGCTCGGACGTACGCGCCGCTGGCGGAAGT 557
130 L...AsnPheGlyValIle...AspGluArgLeuHis..ArgAsnArgG 144
558 GATCAAAACACGCTGTCAATTCGATGCCGAATTTGTTGACGCGCGAA 607
144 uTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAsp 161
608 CACTCGC.....CGCAACCCAAACGCTGCGACGCGGACCAATCCG 645
161 LuTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGlu 177
646 GCTACCTGGGATGCGACCGATCCGCGCTCACCGGTACAGACCA... 692
178 ProTrilEhshslaprogInlycscGlyAsnSerSerArgThrI 194
693 .....CATCAACGCGCGCGCTGTGCGTCAAAATGACAGCTTGAT 733
194 eThr 195
734 CACG 737

seq_name: /cgn2_6/ptodata/1/pna/us60_NEW_COMB.seq:us-60-360-039-32252

seq_documentation_block:
; Sequence 32252, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 32252
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Ralstonia metallidurans
US-60-360-039-32252

alignment_scores:
Quality: 82.50 Length: 313
Ratio: 0.640 Gaps: 16
Percent Similarity: 41.214 Percent Identity: 20.447

Alignment_block:
US-09-528-682-1 x US-60-360-039-32252 ..
Align seg 1/1 to: US-60-360-039-32252 from: 1 to: 1524

8 AlaAspSerArgProProAspGluIleLysArgSerGly..... 20
415 GCGGACGCGCTCGGATGAGAGAGCGCTTACCGTACGCGCAGCACTGCG 464
21 GlyLeuMetProArgGlyHisAsnGlyTyrPheAspArgGlyThrGln 37

```



```

465 GGGGTTTCTACACCGGAGCACCACCGGCTTCCCGAAGGGGGATGC 514
37 eLAsnIleasnLeuTyraSpHisAlaArgGlyThrGlnThrGlyPheVal 53
515 TCAGC.....CATGGAACTGATGCTCTGGCATGCAAC 549
54 ArgTyraSpAspGlyTyraValSerThrSerLeu..... 64
550 CAGCGGATGTCGGGACGCTGATGAATCCATCGATGATGATGAT 599
65 .....SerLeuArgSerAlaHisLeuA 72
600 GCCGATGTTCCACCTGCGATGCTTCCGCTCGATCAACCGTGTCTCG 649
72 IagIyGlnSer.....IleLeu 77
650 TCGGTGGCAGCATGTCGCGATATCCGCTTCGATCCTGCCCGCATGATG 699
78 SerGlyTyraSerThrTyraIleTyraValIleAlaThrAlaProAsnMe 94
700 GAGGGGATTTGGCAGACAGGATGACGGCGCTCGCTGGCGCGCATCAT 749
94 tPheAsnValAsnAspValIleuGlyValTyraSerProHisPro..... 108
750 GATCCAGATGGGG.....CTGGACTGGATGACCGCGCATCCGGGGCGCG 793
109 ..TyraGlnGlnIleValSerAlaLeuGlyIleProTyra..... 121
794 CGCGCGAGCTGGACCTGTGCTGGACAGACCTGGCTATAGCGGCTCG 843
121 ..... 121
844 CCGATGACCCCTGGCGCTCTGGAGGCGACAGGACAGACTTCCAGCGGT 893
122 .....SerGlnIleTyraGlyTrp..... 127
894 GCGCTTTTCGAGGGCTATGGCATGACCGAGCTGGACCGGTGGCCACA 943
127 ..... 127
944 TGCITGGCCCGGAGTATACAGACGCGGCGGCGGCGGCAAGATG 993
128 .....TyraTyraValAsnPheGlyValIleAspGln 137
994 TATTCAGTCGCGCGCGCGCATACGCTGAGGTCAAGTGTGTCGACTTC 1043
137 uArgLeuHisArgAsnArgGluTyraArgAspArgTyraTyraGlnLe 153
1044 CAGGGGCAACAGGTCCGCGGGTACCGTGGCGAGATCATCGTACGTG 1093
153 uAsn.....IleAlaProAlaGlu..... 159
1094 GCCCGAATGTATGCTCGGCTACTGGAACCGCGGACGCTACCGCGAG 1143
160 .....AspGlyTyraArg..... 163
1144 GCGATCCGCAATGGCTGATGATACCGGCGAGCGTGGTACATGACAA 1193
164 .....LeuAlaGlyPheProPro.....AspHisGlnAla 173
1194 GCAGCGCTTCTGTTCTGTGTCGACCGCGCTCAAGACATGATCATCACCG 1243
173 aTTPArgGluGluProTrpIleHisHisAlaProGlnIleCysGlnAsn 190
1244 GTGGCGAAGACT...CTATTCGCGCGAGGTGGAGAGAGCGCTGGCGAGT 1290
190 eTser.....ArgThrIleThrGlyAspThrCys 199
1291 CATCGGCGCGTGGCGAGTGTGTGTATGCGGTG 1325
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1507

```

```

seq_documentation_block:
; Sequence 1507, Application us/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: us/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1507
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: B. fragilis
; US-09-540-209B-1507

alignment_scores:
Quality: 82.50 Length: 219
Ratio: 0.851 Gaps: 18
Percent Similarity: 44.292 Percent Identity: 26.027

alignment_block:
US-09-528-682-1 x US-09-540-209B-1507 ..
Align seg 1/1 to: US-09-540-209B-1507 from: 1 to: 3000

17 LysArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTrpPheAsp 33
609 AAGGCGTTTGGACGGTATCAATCCCGAGACATCCG.....TTCGATGC 652
33 gGlyThrGlnMetAsnIleasnLeuTyraSpHisAlaArgGlyThrGln 50
653 AGGTTTCAAGAAGATCATC...GTCCACTGCCATATACCG..... 689
50 hGlyPheValArgTyraSpAspGlyTyraValSerThrSerLeuSerLeu 66
690 .....TTCAGGGGCTGCAACGGGCT..... 710
67 ArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyraSerThrTy 83
711 CGTTCATATCAC.....TACCAAGATTG 733
83 r.....TyraIleTyraValIleAlaThrAlaPro 92
734 GTAAGCGCGGTGTGCGTAATATATGTTTCGACGATCAGTGGGATGTTCC 783
93 AsnMetPheAsnValAsnAspValIleuGlyValTyraSerProHisProTy 109
784 AATGTATATGAAGATACGATATCTTGGAGACATACGAT.....TA 824
109 rGlnGlnIleValSerAlaLeuGlyGlyIleProTyraSerGlnIleTyrg 126
825 CGCTCTGGCGCTCAAGGAGGTCAAGATGATTGATTGTTCA..... 864
126 LyrTrpTyraValAsnPheGlyValIleAspGlnArgLeuHisArgAsn 142
865 .....AAC 867
143 ArgGluTyraArgAspArgTyraTyraArgAsnLeuAsnIleAlaProAla 159
868 GAAGAGATGCAATCC.....TATCAGAACCG.....AA 895
159 uAspGlyTyraArgLeuAlaGly.....PhePro.....ProA 170
896 CGGGGCGCATGCACTGGCGAGATGAGATTTTCCGTACGGGAGATCACCAG 945
170 spHisGlnAla.....Trp...ArgGluGluProTrpIleHis 181
946 AATTACAGATTGGCTCTTCCAAATGTGTAGGAGAGAGACCAATATTATCAT 995
182 .....HisAlaProGlnIleCysGlyAsn.....SerSerAT 192

```


996 TTCGCCCACTACATGACGACGAGGAGGTGTGTCATGATCGAAGAACG 1045
192 gThrllethrgllyasprhrcysasnlglylthrglnasleuserthr 209
1046 AGCGTATACAGC.....GAAAGCGCAATCTTCTCCACAG 1080
209 leTyr 210
1081 CTTCAC 1085
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2633
seq_documentation_block:
; Sequence 2633, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2633
; LENGTH: 723
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-2633

alignment_scores:

Quality:	82.00	Length:	271
Ratio:	0.678	Gaps:	15
Percent Similarity:	44.649	Percent Identity:	20.295

alignment_block:

US-09-528-682-1 x US-09-540-209B-2633 ..

Align seg 1/1 to: US-09-540-209B-2633 from: 1 to: 723

14 AspGluIleLysArgSerGlyGlyLeuMetProArgGlyHis..... 27
16 GAAAGATATAAATCTGCGTCATCTTTATAGATAAACAACCTCGGCT 65
28AsnGluTyrPheAspArgGlyThrGlnMetAsnIleA 40
66 ACACACAATGAAGAACCAACGCTACTCAGCGCGCAAGAGCTGTCC...A 112
40 snLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAsp 56
113 ATACTTTAGCCATGCGCAGACCCCTGCTGGCATCACCGCA..... 156
57 AspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHis..... 70
157 ...GGATATTTTCTTTGGAAAAGCGCTCGCGCAACCTCACCCCTATTG 203
71LeuAlaGlyGlnSerIleLeuSerG 79
204 GCAACACAGCTGTGTACTGCTTATCTGTGTGGG.....ATGCTGCAT 247
79 LYTyr.....SerThrTyrTyr..... 84
248 CGTACATCATGTTCCAGCTGTATCAGCGCTCCAGCCCGCAAGCGGAAA 297
85IleTyrValIleAlaThrAl 91
298 GAACCTTTCGCGCAATTTGACACAGCGCCCATTTTACCTG..... 336
91 AprOAsnMetPheAsnValAsnAspValLeuGlyValTyrSerPro..Hi 107
337CACATGACAGCGGCACTTACACTCCGTTCA 364
107 sProTyrGluGlnGluValSerAla.....LeuG 117
|||||

365 CCCT.....GTTAGTGTACAGCCATGCCGGGAGTGGGCTGG 402
117 LysGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 133
403 GGAATCTTCACATTGTGTATGCGCTGTGGCTATTTGATGATCATATAGC 452
134 ValIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
453 CTT.....CAAAAACSTGAAGAACACAGCAATCTCGAGACATCTGT 496
150 rArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
497 TTGTAGGATGGGCTCTGCAATCCGTGGTGGCCCTCAACCACTATATGAC 546
167 heProProAspHisGln.....AlaTrpArgGlu 176
547 TCCCTGTCCGATCGGCGGCTCCGCTCTGTGTGCTGTCTGTCTGAGAG 596
177 GluProTyrPheHisAlaProGlnGlyGlyAsnSerSerArgGly 193
597 AGGACCTTCGTATATCAT.....GGAGCGCT 625
193 rIleThrGlyAspThrCysAsnGlyGlyThrGlnAsnLeuSerThrIleT 210
626 TTTACTC.....GTTCGGAACCTTAT 648
210 YrLeuArgGlyTyrGlnSerLysValLysArgGlnIlePheSerArgTyr 226
649 ATGCAGCGCGTTCCTCCACCTGTTTGTCTAGCGGAGAACATAGCCATAT 698
227 GlnSerGluVal 230
699 CATAGCCATCTG 710

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-27746

seq_documentation_block:

; Sequence 27746, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360.039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 27746

; LENGTH: 3291

; TYPE: DNA

; ORGANISM: Neurospora crassa

US-60-360-039-27746

alignment_scores:

Quality:	82.00	Length:	135
Ratio:	1.139	Gaps:	7
Percent Similarity:	53.333	Percent Identity:	25.185

alignment_block:

US-09-528-682-1 x US-60-360-039-27746 ..

Align seg 1/1 to: US-60-360-039-27746 from: 1 to: 3291

56 AspaArgLysTyrValSerThrSerLeuSerLeuArgSerAlaHisLeuAl 72
2017 GACATGGGACAGATCGACGGGAGATTCGACTGAGTGTAACTAGC 2066
72 nGlyIleSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleA 89
|||||


```

353 CCACCGGGCGTGCAGACCAACCAAGGGCCCGCGTTGGGCAAGGGCTCGGCAT 402
38 Asn1IleAsnLeuTyrAspHisAlaArgLYlThrcIlnThrLYlPheValAr 54
   :::::::::::::: ::::: ::::: :::::
403 ACCGTTGGCATGGCGATGGCGGCGCGGTGACAGAGACATTTCAACAA 452
54 gTyraSP.....Asp 58
   |||
453 GCGCCATGCGCGCGCTTCGCATTCAGGAGTACGGCTGTGGCGGCAGC 502
58 lYtYrValSerThrSerLeuSerLeuArGSerAlaHisLeuAlaGlyGln 74
   || :::::::::: ::::: ||| ||||| |||||
503 GCGCATGTGTGAAGAGCATCTCGACAGAGGCGCGCTCGTCGCGGCGCAC 552
75 Ser1IleLeuSer..... 78
   |||||
553 CTGAAGCTGTCCATTCATCTATCTGATCTACAGACAGAACCGGTGAGCAT 602
78 ..... 78
603 TGAAGGCCACACCGACCTCGCTTATAGCGACGAGTGGAAACCGCTTTC 652
79 ..GLTYrSerThrTyrTyrIleTyrVal..... 87
   ||||| ::::: |||||
653 GCGGGTACAACTGCACACGCTGATGTGACGATGCAAGACGCGGCC 702
88 .....IleAlaThrAlaProAs 93
   ||::: ||:::
703 GCGCTGCAAGCGGCATTGTCGGAAGCGCAAAAGCATTTACGACAGCGGCAC 752
93 nMetPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrG 110
   ::::: ||::: ::::: ||:::
753 GCTGATCGTGGTGCACACATCATCGTGTGGGCGCGCCGAT.....A 756
110 lInGlnValSerAlaLeuGlyGlyIleProTyr..... 121
   ::::: ::::: ||| |||
797 AGCAGAGACACCTCGCGCGCGCAGCGCCGCTCGCGCTAGGAAGAGTG 846
122 .....SerGlnIleTyrGlyTyrPyrValAsnPheGlyVal 135
847 GCGCTCGCCAAAGAAAGCATCGCGTGGCGCCGAGATTAATCTCTATGT 896
135 eAspSLuArGLeuHisArg..... 141
   ::::::::::::::
897 GCGGACGGGTGTCAGAACCTTCGCGCGCGCATGCGCGCGCGGCA 946
142 .....AsnArgSLuTyrArGAspArgTyr..... 149
   ::::: ||::: :::::
947 AGCGCGCGCGGAGATGCGACGAGGAGTACAGCGCTACAGCAGAAAG 996
150 TYrArGAsnLeu.....Asn1IleAlaProAlaLysPGLYrArGlye 164
   || ::::: ||::: |||
997 TACCCCGCAACTCGCGGTGAATTGCGCGAGATGCACGCGCAGATTGCC 1046
164 uAlaGlyPheProPAsnHisGlnAlaTyrPArgGluGluProTyrIleH 181
   ||||| ::::: ||| :::::
1047 GCGCGGTGGGACAGCAGCATTTCCACCTTGACAGCGGACCG..... 1089
181 lHisAlaProGlnGlyCysGlyAsnSerSerArgThrIleHisGlyAsp 197
   ||| ::::: ||| :::::
1090 .....AAAGCGTGGTGCACGCGAATGCTGCGGCGAAG 1122
198 ThrCysAsn 200
   ::::: |||
1123 GTGCTCAAT 1131
seq_name: /cgn2_6/prodata/1/pna/US60_NEW_COMB_seq:us-60-360-039-31176
seq_documentation_block:
: Sequence 31176, Application us/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng

```



```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31176
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-60-360-039-31176

```

```

alignment_scores:
  Quality: 81.50      Length: 303
  Ratio: 0.709       Gaps: 11
  Percent Similarity: 37.954   Percent Identity: 19.142

```

alignment_block:

```

US-09-528-682-1 x US-60-360-039-31176  ..
Align seg 1/1  to: US-60-360-039-31176  from: 1  to: 2040

```

```

9 AspSerArgProProasp.....Glu1 16
11:|||||:|||||:|||||
250 GAGCGGACCCGCGAGCCGAGCTTCCTCGACATATCGACMA 299
16 elysargserglyleuwerproarglyhlsanglytyr..... 30
11:|||||:|||||:|||||
300 TTTCCCGCCAGATGGCGACGACAGACACCGGCCCGGAGTACCAGTCA 349
31 .....Pheasparglythrghmet 37
11:|||||:|||||:|||||
350 CCACGCGGTGAGACACACGCGGCCGCTTGCGCAAGGCGTCCGCAAT 399
38 AsnIleasnleutyrsphsialarglythghlthglyphevalar 54
11:|||||:|||||:|||||
400 AGCGTGGGATGGCGATGGCGGCGCTGTAGAGAGCCATTTCACAA 449
54 gtyrasp.....Aspg 58
11:|||||:|||||:|||||
450 GCGCGATGGCGCGCTTCGATATCGGCTTCAGCGCTTGGCGGCGAG 499
58 ltyrvalserthrserleuserleuargseralshlsleualaglyln 74
11:|||||:|||||:|||||
500 GCGACATGATGAAAGCATCTCGCACGAGCGGCTTCGCTCGCGGCGAC 549
75 SerIleuser..... 78
11:|||||:|||||:|||||
550 CTGAGCTGTCCAACTGATCTGATCTACGACAGCAACCGCGTGAGAT 599
78 ..... 78
11:|||||:|||||:|||||
600 TGAAGCCACACGACCTGCTTATAGCAGACGAGTGAACCCGCTTC 649
79 ..glytyrserthrtyrtyrleuval..... 87
11:|||||:|||||:|||||
650 GCGGTGACAACTGGCACACGCTGATGTAGCATGCAACGACGCGCGC 699
88 .....Ilealathrlapras 93
11:|||||:|||||:|||||
700 GCGCTGGAAACGGCATTCGTGAAGCGAAAGCAATTAACGACAGCGCAC 749
93 nmetpheasnvalasnspvalleuvaltyrserprohlsprotyrg 110
11:|||||:|||||:|||||
750 GCTGATCGTGTGACACGATCATCGCTGGGGCGCCCGCAT.....A 793
110 luginleuvalseralaleuuglylyleprotyr..... 121
11:|||||:|||||:|||||
794 AGCAGGACACCTCGCGGCGCACGCGGAGCGCGCTCGGCGTAGAAGAGTG 843

```

```

122 .....SerGlnleuTyRTrpTyRArgValAsnphleuVal1 135
844 GCGCTGCGCCAGAAAGCGTACGGCTGGCCGAGAGATTAATCTTCTATGT 893
135 eaSRgluArgleuHlsArg..... 141
894 GCGGAGCGGTGTGACAGAACGCTTCGGGGGGGCGATCGGGCGGGCGCA 943
142 .....AsnArgGluTyRArgSparGlyTyr..... 149
944 AGCGCGCGCGCGAGAAATGCGACGCGAAGTACGACCGCTACACAGAAAG 993
150 TyRArgAsnleu.....AsnIleAlaProAlaGluAspLytyrArgle 164
994 TACCCGCAACTCGCGCGTGAATTCGCGCAGATCGACAGACGACAAATGCC 1043
164 uAlaGlyPheProProaspHlsAlaTrpArgGluGluProTrpIleH 181
1044 GCGCGGTGGACAGCAGACATTCACACTTCGACGCGGACCG..... 1086
181 lshlsAlaProGlnGlyCysGlyAsnSerSerArgThrIleHrGlyAsp 197
1087 .....AAAGCGGTGGCTGCACGCGAATCGTGGCGCAAG 1119
198 ThrCysasn 200
1120 GTGCTCAAT 1128

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-42281
seq_documentation_block:
; Sequence 42281, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42281
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Halobacterium sp. NRC-1
US-60-360-039-42281

alignment_scores:
  Quality: 81.00      Length: 127
  Ratio: 1.397       Gaps: 5
  Percent Similarity: 45.669   Percent Identity: 24.409

```

alignment_block:

```

US-09-528-682-1 x US-60-360-039-42281  ..
Align seg 1/1  to: US-60-360-039-42281  from: 1  to: 2004

```

```

98 AsnAspValleuGlyValTyRserProhlsProTyRluginleuValse 114
11:|||||:|||||:|||||
418 GACGACATCTCGACGCTGATACATGCGATGCGCGGATGCGGATCGC 467
114 lAlaLeuGly.....GlyIleProTyRserGlnIleTyRcGlyT 127
11:|||||:|||||:|||||
468 GATGCTGCGCTGCGCGCGCTCGCGCGCGCACACTCGGTGTGTGCGG 517
127 rTyRArgValAsnphleuValIleAspGluArgleuHlsArgAsnArg 143
11:|||||:|||||:|||||
518 GCTTTCGCGG..... 528

```


alignment_scores:

Quality: 81.00 Length: 124
Ratio: 1.209 Gaps: 6
Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:

US-09-528-682-1 x US-10-109-551-1 ..

Align seg 1/1 to: US-10-109-551-1 from: 1 to: 78056

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgVa 130
      :::::::::::::::::::: ||| |||
65969 GCAGTGGTAGGGGCGCTT.....GGTGGCTACATGCT 66000
      ::::::::::::::::::::
130 LaspNheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArgVa 147
      :::::::::::::::::::: ||| ::::::::::::::
66001 GGGAAGTGCATAGCAGCGCCCTCTTATACATTTTGGCAGTGCATATGAGC 66050
      :::::::::::::::::::: ||| ::::::::::::::
147 spArGlyTyrArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
      :::::::::::::::::::: ||| ::::::::::::::
66051 ACCGTTACTTCTGTGAACATGACCGCTTACCCCAACCAAGTGTACTAC 66100
      :::::::::::::::::::: ||| ::::::::::::::
163 ArgLeuAlaGlyPheProAspHisGlnAlaTTrparGluGluProTyr 179
      :::::::::::::::::::: ||| ::::::::::::::
66101 AGG.....CCAGTGGATCAGTATAGTAAACCAAGAACACTT 66135
      :::::::::::::::::::: ||| ::::::::::::::
179 pIleHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
      :::::::::::::::::::: ||| ::::::::::::::
66136 TGTGCAAT.....GACTGTGTCAACATCAGTCAGCAAGCAACACA 66173
      :::::::::::::::::::: ||| ::::::::::::::
193 hTleThrGlyAspThrCysAsnGluGlnThrGlnAsnLeuSerThrIle 209
      :::::::::::::::::::: ||| ::::::::::::::
66174 CAGTCACACACACACACAGGGGAG.....AACTTCACCGAAACT 66214
      :::::::::::::::::::: ||| ::::::::::::::
210 TyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTyr 226
      :::::::::::::::::::: ||| ::::::::::::::
66215 GACATCAAGATGATGGAGCGAGTGGTGCACAATGTGCAATTACCCAGTA 66264
      :::::::::::::::::::: ||| ::::::::::::::
226 rGlnSerGluValAspIleTyr 233
      :::::::::::::::::::: ||| ::::::::::::::
66265 CCAGAGAGAAATCCACGCGCTTAT 66286
      :::::::::::::::::::: ||| ::::::::::::::

```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-831-271A-7

seq_documentation_block:

```

; Sequence 7, Application US/09831271A
; GENERAL INFORMATION:
; APPLICANT: RIESMEIER et al., Jorg
; TITLE OF INVENTION: PROMOTERS FOR GENE EXPRESSION IN ROOTS OF PLANTS
; FILE REFERENCE: 0147-0225P
; CURRENT APPLICATION NUMBER: US/09/831,271A
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (256)..(1197)
US-09-831-271A-7

```

alignment_scores:

Quality: 80.50 Length: 139
Ratio: 1.045 Gaps: 7
Percent Similarity: 55.396 Percent Identity: 22.302

alignment_block:

US-09-528-682-1 x US-09-831-271A-7 ..

Align seg 1/1 to: US-09-831-271A-7 from: 1 to: 1420

```

104 TyrSerProHis...ProTyrGluGlnGluValSerAlaLeuGlyGlyI 119
      :::::::::::::::::::: ||| ::::::::::::::
123 TATTCACCTTACAAATCTTATGAATCAATCAGACTCAACATATATAAAGT 172
      :::::::::::::::::::: ||| ::::::::::::::
119 eProTyrSerGlnIleTyrGlyTyrPyrArgValAspNheGlyValIleA 136
      :::::::::::::::::::: ||| ::::::::::::::
173 ACCAACCACAGTAGTCAAAAGT.....GAAGACTTCAAGTACCCT 213
      :::::::::::::::::::: ||| ::::::::::::::
136 spGluArgLeuHisArgAsnArgGluTyr.....Arg 146
      :::::::::::::::::::: ||| ::::::::::::::
214 CAGAG.....TCGAAAGAGATATAGTCGTCAATTTTTCGCAAA 254
      :::::::::::::::::::: ||| ::::::::::::::
147 AspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrAr 163
      :::::::::::::::::::: ||| ::::::::::::::
255 AATGATTACTATAGAAGCCATCAAT.....TCAAGAGTAACTATAA 298
      :::::::::::::::::::: ||| ::::::::::::::
163 gLeuAlaGlyPheProAspHisGlnAlaTTrparGluGluProTyrPi 180
      :::::::::::::::::::: ||| ::::::::::::::
299 GAAAGTATCATTTGTCCGACATGATCATTTCTGCCAAAGATGACT 348
      :::::::::::::::::::: ||| ::::::::::::::
180 LeHisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGly 196
      :::::::::::::::::::: ||| ::::::::::::::
349 ACTACAGAAGCCA.....TTATTTTCGAA 374
      :::::::::::::::::::: ||| ::::::::::::::
197 AspThrCysAsnGluGlnThrGlnAsnLeuSerThrIleTyrLeuArgI 213
      :::::::::::::::::::: ||| ::::::::::::::
375 GATTAACATCAAGAAGAGATCA.....TATGTTCAAGA 406
      :::::::::::::::::::: ||| ::::::::::::::
213 uTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGlnSerGluV 230
      :::::::::::::::::::: ||| ::::::::::::::
407 GGTACCTCGAAGCTTAACCAAGATATTAAGAGACTCATTTTTCAAAT 456
      :::::::::::::::::::: ||| ::::::::::::::
230 AlaSpIleTyrAsnArg 235
      :::::::::::::::::::: ||| ::::::::::::::
457 TTGACTACTTCAAGAG 473
      :::::::::::::::::::: ||| ::::::::::::::

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-9

seq_documentation_block:

```

; Sequence 9, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; FILE REFERENCE: TAMK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Cervus elaphus
US-10-109-551-9

```

alignment_scores:

Quality: 80.00 Length: 124
Ratio: 1.194 Gaps: 6
Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:

US-09-528-682-1 x US-10-109-551-9 ..

Align seg 1/1 to: US-10-109-551-9 from: 1 to: 830

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgVa 130
      :::::::::::::::::::: ||| |||
397 GCAGTGGTAGGGGCGCTC.....GGTGGCTACTTGTCT 428

```



```

130  LasnPhcglyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
429  GGGAGTGCATGAGCAGCGCTTATCATTTTGCAATGACATATGAGG 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147  SPATGYTYTYArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
479  ACCGTTACTATCGTAAACATGTACCTTACCCCAACCAAGTGATAC 528
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163  ArgLeuAlaGlyPheProAspHisGlnAlaIleTyrArgGluIleProTr 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
529  AGG.....CCAGTGCATCATATATATTAACCAAGAACCTT 563
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179  PILEHISAlaIleProGlnGlyCysGlyAsn.....SerSerArgT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
564  TGTGCAT.....GACTGTGCATCATCATCAGTCAAGCAACACA 601
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193  hrlleThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602  CAGTCACCCACCACCCCAAGGGGAG.....AACTTCACCGAAACT 642
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
210  TyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTy 226
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
643  GACATCAAGATGATGAGCGGCTGTGAGCAATGTGCATCACCCAGTA 692
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226  rGlnSerGluValAspIleTyr 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
693  CCAGAGACATCCGAGGCTTAT 714
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-096-080-25

```

seq_documentation_block:
; Sequence 25, Application US/10096080
; GENERAL INFORMATION:
; APPLICANT: Chesbro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Piola, Susette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/10/096,080
; CURRENT FILING DATE: 2002-03-11
; PRIOR APPLICATION NUMBER: US/09/823,494
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/128,450
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Hamster sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (60)..(824)
US-10-096-080-25

```

alignment_scores:

Quality:	80.00	Length:	124
Ratio:	1.176	Gaps:	6
Percent Similarity:	54.839	Percent Identity:	24.194

alignment_block:

US-09-528-682-1 x US-10-096-080-25 ..

Align seg 1/1 to: US-10-096-080-25 from: 1 to: 1000

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVal 130
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417  GCCGGGAGGGGGGCTT.....GCTGCTACATGCT 448
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

130  LasnPhcglyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
449  GGGAGTGCATGAGCAGCGCCATGATCATTTGGCAATGACTGGGAG 498
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147  SPATGYTYTYArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
499  ACCGTTACTATCGTAAACATGTACCTTACCCCAACCAAGTGATAC 548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163  ArgLeuAlaGlyPheProAspHisGlnAlaIleTyrArgGluIleProTr 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
549  CGG.....CCAGTGCATCATATATATTAACCAAGAACCTT 583
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179  PILEHISAlaIleProGlnGlyCysGlyAsn.....SerSerArgT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
584  TGTGCAT.....GATTGTCAACATCATCACCATCAAGACACACA 621
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193  hrlleThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
622  CAGTCACCCACCACCCCAAGGGGAG.....AACTTCACCGAGACC 662
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
210  TyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTy 226
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
663  GACATCAAGATGATGAGCGGCTGTGAGCAATGTGTACACCCAGTA 712
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226  rGlnSerGluValAspIleTyr 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
713  TCAGAGAGATCCGAGGCTTAC 734
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-144-781-41

```

seq_documentation_block:
; Sequence 41, Application US/10144781
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Gangadharan et al.
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DRUG TARGET
; TITLE OF INVENTION: GENES, PROTEINS ENCODED BY THESE MOUSE GENES, AND USES
; FILE REFERENCE: C1001236
; CURRENT APPLICATION NUMBER: US/10/144,781
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 2701
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-144-781-41

```

alignment_scores:

Quality:	80.00	Length:	209
Ratio:	0.792 <td>Gaps:</td> <td>9 </td>	Gaps:	9
Percent Similarity:	48.325	Percent Identity:	20.574

alignment_block:

US-09-528-682-1 x US-10-144-781-41/rev ..

Align seg 1/1 to reverse of: US-10-144-781-41 from: 1 to: 2701

```

12  ProProAspGluIleLysArgSerGlyGly.....LeuMetProAr 25
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2380  CCTACTGATGAACTGTGACAGGTGTGAGCGCGCAATTTTGGGGCA 2331
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25  gGlyHisAsnGluTyrPheAspArg.....GlyThrGln 37
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2330  GCCCTATGCCGTAGCCCTGATGCGCAAAAGGTTTCCACGAGTGAGCAGC 2281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37  etAsnIleAsnLeuTyrAspHisAla..... 45
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2280  TTGCATCCCGCATCTATGACACCTCATATCCAGTACGTATTCATCCAT 2231
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
46  ..ArgGlyThrGlnThrGlyPheValArgTyrAspAspIlyTyrValSe 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

2230 GATGAGGCGCTTGAGCTTGGC..... 2210
61 rThSerLeuSerLeuArgSerAlaHisLeuAlaGlyInSerIleLeuS 78
2209 .....GGGTGGCTCCGTCGACATGGCCATCCGTCGTCGA 2176
78 eGcLyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMet 94
2175 GTGGTAGGTGGCGTGTGGCGACGATGTGTGGCGT.....CATGTC 2135
95 PheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGlyL 111
2134 GGGGAGCTCCGCTTGATGTGATGTCCGCGCTCTCTCCACACGCGTGC 2085
111 nGluValSerLeuLeuGlyGlyIleProTyrSerGlnIleTyrGlyTrrT 128
2084 CAAAGCGGAGCCCTTGAGAGGCTG..... 2060
128 yTArGValAsnRheGlyValIleAspGluAlaGlyLeuHisArgAsnArgGlu 144
2059 .....GTCCACCTTGCGGATCATGGATTCGACA..... 2033
145 TyTArGAspArgTyrTyTArGAsnLeuAsn...IleAlaProAlaGluAs 160
2032 .....CAGCTCTCTCAATGTCTTGTCCGCGACATGCA 2001
160 pGcLyTyrTArGleuAlaGlyRheProProAspHisGlnAlaTrrArgGlu 177
2000 CAGCTGGCAGGTGGCTGTGTAACTGATGACGATGACGACGACGAGAAAGTT 1951
177 LuProTrrPrlHisHisAlaProGlnGlyCysGlyAsnSerSerArgThr 193
1950 GCCCAGAGATTCATGAG.....GAGCGGTCCGCTAGACACACTGGGTGT 1907
194 lIethrGlyAspThrCysAsnGluGlu 202
1906 CTTACTGGAGACAGTGGCTCCGAAAGAG 1880

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2740

seq_documentation_block:
; Sequence 2740, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2740
; LENGTH: 3264
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-2740

alignment_scores:
Quality: 80.00 Length: 219
Ratio: 0.748 Gaps: 12
Percent Similarity: 48.858 Percent Identity: 19.635

alignment_block:
US-09-528-682-1 x US-09-540-209B-2740 ..
Align seg 1/1 to: US-09-540-209B-2740 from: 1 to: 3264

6 TyTArGAlaAspSerArgProAspRplUleUlysArgSerGlyGlyLe 22
1558 TACACGCGAGTAACGAA...CCGACCTGATTCATGACAGCGGCGGC 1604
22 uMeTProArGlyHis...AsnGlyTyrPheAspArgGlyThrGlnMet 38
::: ::::: ||||| ::: :::

```

```

1605 ACTCGGCTCTCTCAGTAACGAGTATTCATATGAAGAAACCTAGTACA 1654
38 snlIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
::: ||||| ::: ||| :::
1655 GCAGCAGCATCTACTCGGACTAC.....TTTAAGAG 1686
55 TyTArAspArgLyTyrValSerThrSerLeuSerLeuAlaGSerAlaHisLe 71
1687 TTTCACACGCGACACTACTCAAACTTATGGCGCGTTTCACACAGTACST 1736
71 uAlaGlyInSerIleLeuSerGlyTyrSerThrTyrTyrIle..... 85
1737 T.....TATTAACCCGTTATGTGTCGACGCTC 1762
86 .....TyTValIleAlaThrAlaProAsnMetPheAsnValAsn 98
1763 AAAGAGCAGACTGATTCAGATTGCGGTCCATCCATC..... 1800
99 AspValLeuGlyValTyrSerProHisProTyrGlnGlnGluValSerAl 115
1801 .....AACACAGCCACCGAAGACCCGAAAGC 1826
115 AlAuGlyGlyIleProTyrSerGlnIleTyrGlyTrrTyr...ArgValA 131
1827 ATGGGAGGAGATATGCGCACAAATGCGGTAGCGGTTTCTTGCGAGTATCA 1876
131 snPheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrArgR 147
1877 ACTAC.....AATTATTAAGAC 1893
148 ArgTyTyrTArGAsnLeuAsnIleAlaProAlaGluAspGlyTyTArGle 164
1894 CGATCAGATGGAGAGTAACTAAC...GGACGTTACGACGTTCTACGCTT 1940
164 uAlaGly.....PheProProAspHisGlnAlaTrrA 175
1941 TATAGTGAACAAGCCTGGGATCTCTCCCTTCATTCCTGCGCGAGTGA 1990
175 rg.....GluGluProTrrPrlHisHisAlaProGlnGlyCysGlyAsn 189
1991 ACGTTGCACAGAACCATTCCTTGAAGAAATCCGCGAAAGACGATATC 2040
190 SerSer.....ArgThrIleThrGlyAspThrCysAsnGlnGluThr 203
2041 GGTACATTAAGAACTTCGCGCTTCTTGCGGACAAATTAAGTAATACAGATAC 2090
203 rGlnAsn 205
2091 CAAAGAC 2097

seq_name: /cgn2_6/ptodata/1/pna/PCT_PCT_NEW_COMB.seq:PCT-US02-09921-59

seq_documentation_block:
; Sequence 59, Application PC/TUS0209921
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DAFFO, Abel
; APPLICANT: JONES, Anissa L.
; APPLICANT: TRAN, Alanna-Phung B.
; APPLICANT: DAHL, Christopher R.
; APPLICANT: GIETZEN, Darryl
; APPLICANT: CHINN, Joyce
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: TUASON, Olivia
; APPLICANT: VAP, Pierre E.
; APPLICANT: AMSHEY, Stefan R.
; APPLICANT: DAUGHERTY, Sean C.
; APPLICANT: DAM, Tommy C.
; APPLICANT: LIU, Tammy F.
; APPLICANT: NGUYEN, Duy-Viet An
; APPLICANT: KLEEFELD, Yael

```



```

167 heProboAphisgInlaItrpAgluProTyr..... 179
168 ::::::::::::::::::::::::::::
169 ATCAAGACAGGAGCAATCTGCGCGGACAGATATAAAGAACAGAC 469
170 .....1leHisHislaProgInglCysGlyAsnSerSer 192
171 ::::: ||||| ::::: |||||
172 CGGCGGAAGTGCCTCATCCCATCCATCCAGGCAATCACAGATCAGCG 519
173 ghrIlethrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThr 209
174 ::::: ||| ::::: ||::: ||::: ||:::
175 gTCrATCATGTGCGATTAGCA.....GAAGCAAGCGCTTGCTC 563
176
177 209 letYrleuAargLutYrGlnSerLysValLysArg 220
178 ::::: ||| ::::: ||::: ||::: ||:::
179 TCCTGCTTTCAGTTCCAGTTCCAGTTCTTATTCAGAAA 598
180
181 seq_name: /cgn2_6/ptdata/1/pna/US60_NEW_COMB.seq:US-60-360-039-26953
182
183 seq_documentation_block:
184 ; Sequence 26953, Application US/60360039
185 ; GENERAL INFORMATION:
186 ; APPLICANT: Cao, Yongwei
187 ; APPLICANT: Chen, Xianfeng
188 ; APPLICANT: Goldman, Barry S.
189 ; APPLICANT: Hinkle, Gregory J.
190 ; APPLICANT: Slater, Steven C.
191 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
192 ; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
193 ; FILE REFERENCE: 38-10(52052)A
194 ; CURRENT APPLICATION NUMBER: US/60/360,039
195 ; CURRENT FILING DATE: 2002-02-21
196 ; NUMBER OF SEQ ID NOS: 47374
197 ; SEQ ID NO 26953
198 ; LENGTH: 769
199 ; TYPE: DNA
200 ; ORGANISM: Neurospora crassa
201 US-60-360-039-26953
202
203 alignment_scores:
204 Quality: 79.00 Length: 267
205 Ratio: 0.731 Gaps: 13
206 Percent Similarity: 40.449 Percent Identity: 20.599
207
208 alignment_block:
209 US-09-528-682-1 x US-60-360-039-26953 ..
210
211 Align seg 1/1 to: US-60-360-039-26953 from: 1 to: 769
212
213 1 AsnGlyAspArgLeuTyrArg.....AlaAs 9
214 ::::: ||||| ::::: |||
215 CACGGGCAAGGACTTTACCATCATGTGCGCGTCCAAAGCCGCTATGCGG 64
216
217 9 pSerArgProPaspGluIleLysArgSerGlyGlyLeuMetProArg 26
218 ||| ::::: ||::: ||::: ||:::
219 GAGCGACCATCTCTGAAGCATCCGACGACAAGACAAGATCGCTGAACAAG 114
220
221 26 LysHisAsnGlu..... 29
222 ::::: ||::: ||::: ||::: ||:::
223 CACACACTGATGTGCATTCTCCGGTGAAAGCTGCGACACAGGTACGAGC 164
224
225 29 ..... 29
226
227 165 GATATCGATGATGCGCGCTTTGACCAACAACCAAGTTAACCAACTGTT 214
228
229 30 .....TyrPheAspArgGlyThrGlnMetAsnIleAsnL 41
230 ||| ::::: ||||| ||| :::::
231 GTTCCCATCGACAGTGAAGTTGCGCGAGTACACCAAGCAAGCCAGC 264
232
233 41 euTyrAspHisAlaArgGlyThrGln.....ThrGly 51
234 ||||| ::::: ||||| ::::: ||::: ||:::
235 TCTATCTATGCGCAAGCGGACCGACTTTCCTCCCTTCCCTTGCAC 314

```



```

52 PheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeuArgSer 68
   |||||
315 TTGTGCGC.....GGCAGCTTGCGAGAGCTTGCGCTCCGCCAA 355
68 rAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrI 85
   :
356 ACCC..... 359
85 leryValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeu 101
   :|||
360 .....TACAACGTAAACCTCTCTC 380
102 GlyValTyrSerProHisProTyrGlnGln..... 112
   |||
381 GGAAGGTGCACCCCTATCACCCACAGCCAGCTTACTGCTGCAGTA 430
113 .ValSerAlaLeuGlyGlyLeuProTyr...SerGlnIleTyrGlyTyr 128
   :|||
431 CTTGGGCTGTTGGCGAAGGTCCCTATGCGCGCCAGGCTACGACAGT 480
128 yArgValAsnPheGlyValIleAspGlnArgLeuHisArgAsnArgGln 144
   |||
481 ATGTGATCTGCACCGATCTGTACCGCTTGTATGCGTAGAGTCAAT 530
145 .TyrArg.....AspArgTyrTyrArgAsnLeuAsnIleAlaProAla 159
   |||||
531 GTACAGATCACTGACCGGTTCTTCTCTCATCAGATCACTGCGCTCG 580
159 LuAsP.....GlyTyrArg 163
   :|||
581 ATTCTGCACAGCAGCAGCAGCAGATCAGCTCCACAGAGGTATCA 630
164 LeuAlaGlyPheProProAspHisGlnAlaTyrArgGlnLeuProTyrI 180
   :|||
631 GCTTCT....CAACCTGTGCACCGAGAGCTCAAGAGAG.....GT 668
180 eHisHisAlaProGlnGly.....CysGlyAsnSerSerArg 192
   :|||
669 TACCATAGACTTCAGGCGATGACGCTCAAGGCGTACAGAGAGCGG 717
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-7

seq_documentation_block:
; Sequence 7, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAKK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Odocolleus hemionus hemionus
US-10-109-551-7

```

```

alignment_scores:
  Quality: 79.00      Length: 124
  Ratio: 1.179      Gaps: 6
  Percent Similarity: 54.032      Percent Identity: 24.194

```

```

alignment_block:
US-09-528-682-1 x US-10-109-551-7  ..
Align seg 1/1  to: US-10-109-551-7  from: 1  to: 771

```

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVal 130
   |||||
367 GCAGTGTAGGGGCTC.....GGTGGTACATGT 398
130 lAsnPheGlyValIleAspGlnArgLeuHisArgAsnArgGlnTyrArg 147
   :|||
399 GCGAAGTCCATGAGCAGGCGCTTATACATTTTGGCATATCATGAGG 448
449 ACCGTACTACTCGGAACATGATGACCGTTACCCCAACAGTGTACTAC 498
163 ArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGlnLeuProT 179
   |||
499 AGG.....CCAGTGCATCAGATATATACCGAAGACCTT 533
179 rIleHisAlaLarProGlnGlyCysGlyAsn.....SerSerArg 193
   :|||
534 TGTGCAT.....GACTGTGCACATCAGATCAGTCAAGACACA 571
193 hIleThrGlyAspThrCysAsnGlnGlnTyrGlnAsnLeuSerThrIle 209
   :|||
572 CAGTACACACACACACACAGGAG.....AACTCACCGAACT 612
210 TyrLeuArgGlyTyrGlnSerIleValIleArgGlnIlePheSerArg 226
   :|||
613 GACATCAAGATGATGAGCGAGCTTGTGAGCAATGTGCATCAGCCAGTA 662
226 rGlnSerGluValAspIleTyr 233
   |||
663 CCAGAGAGATCCAGGCTTAT 684
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-5

seq_documentation_block:
; Sequence 5, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAKK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Odocolleus virginianus
US-10-109-551-5

```

```

alignment_scores:
  Quality: 79.00      Length: 124
  Ratio: 1.179      Gaps: 6
  Percent Similarity: 54.032      Percent Identity: 24.194

```

```

alignment_block:
US-09-528-682-1 x US-10-109-551-5  ..
Align seg 1/1  to: US-10-109-551-5  from: 1  to: 830

```

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVal 130
   :|||
397 GCAGTGTAGGGGCTT.....GGTGGTACATGT 428
130 lAsnPheGlyValIleAspGlnArgLeuHisArgAsnArgGlnTyrArg 147
   :|||
429 GCGAAGTCCATGAGCAGGCGCTTATACATTTTGGCATATCATGAGG 478

```



```

147 spargTYTfYArg...AsnleuasnleialaProalaGluaspolyTyr 162
|||||
479 ACCGTACTATCGTAAGAAACATGACCTTACCCACCAAGTAGTACTAC 528
163 ArgleualaglyPheProAspHisGlnalaTrrparglugluProtr 179
|||
529 AGG.....CCAGTGCATCGATATATAACCAAGAACACCTT 563
179 PLeHisHisalaproGlnGlyCysGlyAsp.....Serseragt 193
|||||
564 TGTGCAT.....GACTGTGCACATTCACCTCAAGCAACA 601
193 hrleThrGlyAspThrcysasnugluThrglnAsnleuserThrile 209
|||||
602 CAGTCACACACACACCAAGGGGAG.....AACTCACCAGAACT 642
210 TyrleuAtggluTyrGlnSerLysValysArglnIlePheSeraPty 226
|||||
643 GACATTAAGATGATGACGAGTGTGTGAGCAATGTGCATACCCACTA 692
226 rGlnSerGluValaspIleTyr 233
|||||
693 CACAGAGAAATCCAGGCTTAT 714
seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:US-10-106-698-580

```

```

seq documentation block:
; Sequence 580, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 580
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-580

```

```

alignment_scores:
  Quality: 79.00      Length: 178
  Ratio: 1.013      Gaps: 6
  Percent Similarity: 43.820      Percent Identity: 19.663

```

alignment_block:

US-09-528-682-1 x US-10-106-698-580/rev ..

Align seg 1/1 to reverse of: US-10-106-698-580 from: 1 to: 982

```

83 TyrTyrIleTyrValIlealThrAlaProAsnMetPheasnValasnas 99
|||||
955 TTTTGTCTGTAATCTATTCAGTAGATCTTTGCCGAGGCAATCACAATTC 906
99 pValleuglyValtyrSerProHis..... 107
|||||
905 GCTCTCGGAATCTATTCAGTAGATCTTTGCCGAGGCAATCACAATTC 856
108 .....ProTyrGluGlnGluValSerAlaLeu 116
|||||
855 ACACAGGGGTGTGACACGACACTCCCTCGCTACCTGCGCTGTGGCAGGG 806
117 GlyGlyIleProTyrSerGlnIleTyrGlyTrrPtyrArgValasnpheG1 133

```

```

805 GCAGAGAGTGAATGGCTCT.....TCCTGGGCAACAGGCCACAGTTAA 765
133 YValIleaspGluArgleuHisArgAsnArgGluTyrArg..... 146
764 CCGGTGACAAATGCGAGAGCCATAGGAATTCATCTTTAGAAAAAACGAA 715
147 .....AsparGTYTfYArgAsnleuasnleialaProalaGluasp 160
|||||
714 ACTGATTAATAAAAAAACAACCAAACTGAGTATTAACCTCAGCAGTGT 665
161 GlyTyrArgleualaglyPhe..... 167
664 TCCAGCACTATCTCGGGCTTAAATATAAAGGCATCATAGAAAAACAG 615
168 .....ProProAspHisGlnalaTrr. 174
614 TTAATAAGATTAACAGACAGAGCCGCCACCTCCAGGCACTGTGATGG 565
175 .....ArgGluGluProTrrPleHisHisalaproGlnGly 186
564 GGTGGGGCCAGCCCCGCTGAGCGCGGGGCTCCTCATTGAGACGAGG 515
187 Cys.....GlyAsnSerSerArgThr11 194
|||||
514 TGCAGGCCCGCGCTGAGCAGCGGCGAGCAGCAGCGCCAGCAGACAGCAG 465
194 eThrGlyAspThrcysasnugluThrglnAsn 205
464 CGCCAGCAGCAGCTGTTCGAGGAGCACCAGGAC 431
seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12700

```

```

seq documentation block:
; Sequence 12700, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12700
; LENGTH: 6078
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12700

```

```

alignment_scores:
  Quality: 79.00      Length: 285
  Ratio: 0.530      Gaps: 15
  Percent Similarity: 52.281      Percent Identity: 21.053

```

alignment_block:

US-09-528-682-1 x US-10-105-299-12700/rev ..

Align seg 1/1 to reverse of: US-10-105-299-12700 from: 1 to: 6078

```

7 ArgAlaSpSerArgProProAspGluIleLysArgSerGlyGlyLeu 23
|||||
1723 CAAGCAATAAAGAGATACAAACAAATGGAAGAACATTCATCGCAT 1674
23 tProArgGlyHisasnugluTyrPheAspArgGly.....ThrGlnMeta 38
|||||
1673 GGGTAA...AGCAATCAATATCGTGAATAGGCCATATCGCCAGAGTA 1627
38 snleuasnleuTyrAspHisalarglyThrglnThrglyPheVal... 53
|||||
1626 ATTTAAGATTCAATGTCATCCCATCAAGCTACCAATGCTTCTTTCAC 1577

```



```

seq_documentation_block:
; Sequence 24895, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24895
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Methanobacterium thermoautotrophicum
US-60-360-039-24895

```

```

alignment_scores:
Quality: 78.50 Length: 175
Ratio: 1.061 Gaps: 11
Percent Similarity: 42.286 Percent Identity: 24.000

```

alignment_block:

us-09-528-682-1 x US-60-360-039-24895/rev ..

Align seg 1/1 to reverse of: US-60-360-039-24895 from: 1 to: 1272

```

85 ILETVALLIALLAATHRALAProsmnerPheasnValasnpValle 101
:::|||||::: ||| |||:::|||||::: ||
585 CTCATGCGAAGGTGGATCCCT.....TTCAAGCATATCCGAGCCTCT 542
101 uGIValIyIser.....ProHisProt 109
|||||::: ||| |||:::|||||::: ||
541 GGGGCTCCCGCAACATCATCAATTCATGCCCTGAGCGCATCTGTG 492
109 YRGUGINGUVALSERALALEUCLYGLIIEPOTYRSEGINILETYR 125
:::|||||::: ||| |||:::|||||::: ||
491 GTCAGTAT.....GTAAATGGCGGTACCGTTTATATTCATATAT 451
126 GLYTPRYARVALASNPHE...GLYVALILEASPGIARGLUHNISAR 141
|||||::: ||| |||:::|||||::: ||
450 GGATTCATATCTGTGGCTGCGAGGAGACTGCTGAG..... 415
141 GASNARGGLUTYRARGSPARGTYRTYRARGASNLEUASNILEALAPROA 158
|||||::: ||| |||:::|||||::: ||
414 .....GGCC 411
158 IAGIUAASPGIYTYRARGLEUALGLYPHEPROPROASPHISGLNALTIP 174
|||||::: ||| |||:::|||||::: ||
410 CTGAGAGCGGGGTATAC.....CACACACCAACTGG 380
175 .....ARGGLUIPROTPRIIEHIS.HIS.....A 183
|||||::: ||| |||:::|||||::: ||
379 CCTCAATGCAACCCATCGCATGCTCCGAGCATCATATATGTCA 330
183 LAPROGLINGLYCYSGLYASNSERSEARGLTHRIETRGLYASPTHCYS 199
|||||::: ||| |||:::|||||::: ||
329 TCCGAGTGCCTTGC.....AAGGAGTGC 307
200 ASNGIUGLUTHIRGINASNLEUSERTHRIETRYLEUARGGLUTYRGINSE 216
|||||::: ||| |||:::|||||::: ||
306 CTGTGTGGCTGTACAAACATGTTTACCATTTTCTAT...GGCCTTCAGGA 260
216 RLYEVALLYSARGLINLEPHESEARGLYRGINSERGLUVALASPILET 233
|||||::: ||| |||:::|||||::: ||
259 TGAAGCTACAGCGCGGTTCATAGCCACCTATGAGTCTTATAGACATGCT 210
233 YRASNARGIIEARGASPLUDEU 240

```

```

209 ATTTCAGATCTTCGAGATGTC 187

```

seq_name: /cgn2_6/pdata/1/pna/US60_NEW_COMB.seq:US-60-360-039-34717

```

seq_documentation_block:
; Sequence 34717, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 34717
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Ferroplasma acidarmanus
US-60-360-039-34717

```

```

alignment_scores:
Quality: 78.50 Length: 263
Ratio: 0.649 Gaps: 12
Percent Similarity: 46.008 Percent Identity: 22.053

```

alignment_block:

us-09-528-682-1 x US-60-360-039-34717 ..

Align seg 1/1 to: US-60-360-039-34717 from: 1 to: 1299

```

16 ILEYSARGSERGLYLUEMETPROARGLYHNISNGLUTYRPHAS 32
|||||::: ||| |||:::|||||::: ||
322 ATAAAGAAATCTGTGGTTCATTGAAG.....CTAAGATTTCAAA 365
32 PARGGLYTHGLMETASNLEASNLEUTYRASPRIASALARGLYTHG 49
|||||::: ||| |||:::|||||::: ||
366 TTCGGGAGTGAAGCCACATGCATGCCATAGCGTCGCAAGGATTC. 414
49 IINTHYR.....PHEVALARGTPASPGIYTYRVALISERTHR 62
|||||::: ||| |||:::|||||::: ||
415 .ACGGGGAAGAAAGTACATTGTAAAAATGGAAGCGGATAC..... 453
63 SERLEUSERLEUARGSERALANISLEUALGLYGLNSETIIELEU..... 77
|||||::: ||| |||:::|||||::: ||
454 .....CACGGTGCACATGACTATGCTCGATTAAC 482
78 .....SERGLYTYRSETHRTYRTYRIETRYVALILEALITHALAP 92
|||||::: ||| |||:::|||||::: ||
483 ATCCGGCAGTGGCACCATGACATTCGAGTTCATCTTCGACGGGTAC 532
92 ROASNMETPHEASNVALASNPVALLEUCLYVALTYRSEPROHISPRO 108
|||||::: ||| |||:::|||||::: ||
533 CGGAGGAAGTTTCCAGAACAGTGTGTTGGAGATACAAATGACCGTGA 582
109 .....TYRGUGINGUVALSERALALEUCLYGLIIEPT 120
|||||::: ||| |||:::|||||::: ||
583 AGCATTTGAAGAAGCTTTCAAGAAACACAGAAATGAATTCCTCGGTAT 632
120 OTRYSERGLIIEYRGLYTYRTRYARVALASNPHEGLYVALILE... 135
|||||::: ||| |||:::|||||::: ||
633 TACAGAACCCGTTCTTGA.....AATATCGGTATATAAATC 670
136 .....ASPGIARGLEHNISARGASNARGGLUTYRARGSPARGTYRTY 150
|||||::: ||| |||:::|||||::: ||
671 CGAGAAATGATTTTGTGAACSTCTACGGGAAATGAATGAAATATGTA 720
151 ARGASNLEUASNILEALAPROALAGIUAASPGIYTYRARG..... 163

```


89 AlaThrAlaProAsnMet.....PheAsnValAsnAspValLeuGlyVa 103
|||||::|||::: ||| :: |||::

737 ATTATGGTGCATCATTCCTGAGCTAC 764

[illegible]

APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CONYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPATION
FILE REFERENCE: BG1-128CP
CURRENT APPLICATION NUMBER: US/09/602,777A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932127.2
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932129.9
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: DE 19932226.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932920.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932924.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932928.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932930.3
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932933.8
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932935.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932973.7
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933002.6
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933003.4
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941390.8
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941391.6
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 442
SEQ ID NO 227
LENGTH: 1020
TYPE: DNA
ORGANISM: *Corynebacterium glutamicum*
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(997)
OTHER INFORMATION: KX01932
US-09-602-777A-227

```

alignment_scores:
  Quality: 77.50
  Ratio: 1.845
  Length: 77
  Gaps: 2

```


Percent Similarity: 54.545 Percent Identity: 29.870

alignment_block:
US-09-528-682-1 x US-09-602-777A-227 ..

Align seg 1/1 to: US-09-602-777A-227 from: 1 to: 1020

```
121 TyrSerGlnIleTyrGlyThrPyrArgValAsnPhcGlyValIleAspG1 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
259 TACGCTTATCTGCGGCGGCGGATTTGGCGTGTGGTGTACTCCACG 308
137 uArg...LeuHisArgAsnArgGluTyrArgAsnArgTyrTyrArgAsn1 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
309 GCGGTGGCTGCGTCCGCGGCGACCCCTACATGATGCGACAGGTTCCAC 358
153 euAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
359 TCCTCT.....AGTGGTTGGCTGCC 378
170 AspHisGlnAlaTyrPargGluProTyrPheIleHisAlaProGlnG1 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
379 GATCATCGCATCGCATCGCTGACTACCGACTTGCACCGCACCCATTTTC 428
186 yCysGlyAsnSerSerArgThrIleThrGly 196
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
429 CAGCGCTATCGACGACGCGCTTGGCAGTGT 459
```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW.COMB.seq:US-09-540-209B-1521

seq_documentation_block:

```
; Sequence 1521, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1521
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: B.fragilis
US-09-540-209B-1521
```

alignment_scores:

Quality:	77.50	Length:	217
Ratio:	0.891	Gaps:	10
Percent Similarity:	40.092	Percent Identity:	19.355

alignment_block:

US-09-528-682-1 x US-09-540-209B-1521 ..

Align seg 1/1 to: US-09-540-209B-1521 from: 1 to: 1293

```
50 ThrGlyPheValArgTyrAsp.....AspGlyTyr 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
658 ACCGGGATCCCGGTATGCTTGATATATACCGTTTGTATGATCTGTA 707
59 rValSerThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSer1 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
708 TTACCAAGGCGGGAATCCCTTGTCCGCCACAGAGGTGGCCAGCCCGTTTA 757
76 leuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaPro 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
758 TCTCACAGCGCGCTTGCACCCCTACATCTTAT..... 789
93 AsnMetPheAsnValAsnAspValLeuGlyValTyrSerProHisProTy 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
790 .....AAGGTATGATCCGATACCTG 812
109 rGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnIleTyrG 126
```

: ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

813 GGCACGATGCTACGCCGGGTCAACGCGCTCACTTCCGCCGAATGTACG 862

126 ly..... 126

863 GGAAACAGGTTGGCATGGCTGGCGCTCCATCAGTTGTCCGACGGGTTTC 912

127 ...TrpTyrArgValAsnPhcGlyValIleAspGlnArgLeuHisArgAs 142

913 ACCTGGAAGAGTACATGACTTCTCTTGACAG.....CTTCC 953

142 nArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaG 159

954 CCGGCGACAGGAGGAACTACTCTGGAGAACTGAGGTTGAC..... 996

159 luAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrArg 175

997CAGAAATCTGAGCA 1011

176 GluGluProTyrPheIleHisAlaProGlnGlyCysGlyAsnSerSer1 192

1012 GAAAA.....GGAGCTGC...CTGGGAGAGGA 1037

192 rThrIle.....ThrGlyAspThrC 199

1038 AACGATCGGAGACCTGCTGCGCGCGGTGTACCGTTACCGTGGAGAAAT 1087

199 yAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGln 215

1088 GC.....ACGACATACCGGACTGAC..... 1107

216 SerLyValysArgGlnIlePheSerAspTyrGlnSerGluValAsp1 232

1108AAGAGCGCCCTCCGATGAGTATATAGACAGATAGATAT 1148

232 e 232

1149 C 1149

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW.COMB.seq:US-09-540-209B-4973

seq_documentation_block:

```
; Sequence 4973, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 4973
; LENGTH: 2847
; TYPE: DNA
; ORGANISM: B.fragilis
US-09-540-209B-4973
```

alignment_scores:

Quality:	77.50	Length:	208
Ratio:	0.881	Gaps:	11
Percent Similarity:	42.308	Percent Identity:	23.077

alignment_block:

US-09-528-682-1 x US-09-540-209B-4973 ..

Align seg 1/1 to: US-09-540-209B-4973 from: 1 to: 2847

```
54 ArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSer..... 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
664 CGCAATAGCGAGAGATATATCACTTCTTCCGCAATCCGCTTCGCGCA 713
69 .....AlaHisLeuAlaGlyGlnSerIle 77
```



```

714 CACGATAGTATGATGAAGACTGAAGCGCCACTACTGTTAATCGCGCGCTAA 763
77 euserglytyrserthrtyrlyleuylvalleala..... 89
764 TGACAAAGACTCTTGTACGTATGAGGCACTATCTCCAAAGCGGCTGAT 813
90 .....Thralaprosme 94
814 GCTTTCAGCAAGAGGACACTTCTTGGAGAGAAATCTGATTAACA 863
94 tpheasvalasnaspyalleucllyvaltyrserprohisprotglyug 111
864 GTTGGAAATTAATACCTTCTGGCGCTGATGTGATATGACAAATTAATA 913
111 lngluvalseralaleuylglylleprotyr.....sergln 123
914 TCTGGACAGCCCTTGACAAATGTAATGCCATATGCCCAATAATTCGCTG 963
124 lletyrglytrp.....tyrargvalasnpheglyvalleaspgluar 138
964 ATTACCATTTTCGAGCGGCTTCGCCGAAGTCGGGATGGTTATGATGT 1013
138 glenuhsararasnargclutyrargasparglytyrtyrargasnleuasn1 155
1014 GCTGGTAACGGGAT.....AAGGATGCTTAC.....A 1039
155 lealproalaglualspglytyrargleu..... 164
1040 TCCGATCCATTCACAGACTGTACCGGCTTAACATACCCCTCGAGCTTA 1089
165 ..Alaglypheproproasphisglnalatrpargsluigluprotpril 180
1090 GTTCCGGGCTT.....GAAGACAGGACATGAGCCATTGGAGAGGTG.. 1131
180 ehshisalarproglngly.....cysglyasnsersearagt 193
1132 .....GGCGGCACAGGTGCTTTCGGTCATTAATAAAGTA 1165
193 hr.....ilethgly 196
1166 CTTTCCAGATTAAAGCATGCACAGCCCGCTTCTTCCGATGTGCGGGG 1215
197 AspThrCysasnlglnluthrgln 204
1216 GCTATGTGTATCGCGCAGCGCGCAG 1239

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41882
seq_documentation_block:
; Sequence 41882, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianteng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10652052/A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 41882
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Thermoplasma acidophilum
US-60-360-039-41882

```

```

alignment_scores:
Quality: 77.00      Length: 260
Ratio: 0.681       Gaps: 12

```

```

Percent Similarity: 43.462      Percent Identity: 21.538
alignment_block:
US-09-528-682-1 x US-60-360-039-41882
Align seg 1/1 to: US-60-360-039-41882 from: 1 to: 1311.

```

```

12 ProProasp1u1leuylsargserglyglyleuemetProargglyhisas 28
388 CGCGCGGAGGACACTACTACTCTCC.....CCGACAAATCCCTC 425
28 nglytyr.....pheasparglythrlnmetasnlla 40
426 CGAGTACGATCTCAACTTCCGCGCTGATGCGGATAGCCATTAACCTTCA 475
40 snleuylrarnhisalar..... 46
476 ACACCTTTCACACAGTACAGAGATAGAGAAAGCCGACAGAGATATCC 525
47 .....glythrlnth 50
526 TTCAGGATCAACCCGGGTTTGGCATGGGGAATTCGACAGTACACAC 575
50 rglyphevalargtyrarnsparglytyr.....valserthrserleus 65
576 CGGGGCTGCCAGGACAAAGTTCGGGATAGACCCGATGCCCATCTCG 625
65 erleuylrserlalnhisleualagly.....glnser 75
626 CCGACAAAGAGCCAGGAGAGACGGGATCAGGAGTCCGCTATACATAG 675
76 lleuenserglytyrserthrtyrlyleuylvallealathralabr 92
676 ATGATCGGATTCGAACAACAGGATCATGTCAAGATCGTGAAGCCATATC 725
92 oasmetpheasnasvalasnaspyalleucllyvaltyrserprohisprot 109
726 GAATTCCTTCGCCATTCGACAGACGATCGGAGGAGGCGGCTGAGT 775
109 yrglytrp.....glnluvalseralaleuylglylleprotyrserglnle 124
776 TTCATATTCGCGGATGTGGGGGCTGGCGCATTCGATATGTCAG... 822
125 tyrglytrptyrargvalasnpheglyvalleaspgluarlguehisar 141
823 ...GGCGAAACGAACTGAGACATTCGGCACATGGGCTGACGCCGTG... 864
141 gsnatrglytyrargasparglytyr..... 150
865 ..TTGAAAGAAATTT..GATCGATACCATTTCCGCGATCTCGCTCGAAC 910
150 ..... 150
911 CGGAGCCATATCTGGTTGGCGATGCCGATCATGCTCGGCACAGTCAC 960
151 .....Argasnleuasn1 155
961 GATGTCCACAGCGGCTTCGGGGCACAGACATCGGATGAACCTCAACAT 1010
155 ealaproalaglualspglytyrargleualaglypheproproasphsg 172
1011 AAGGCGCGGCTGTACGGCGCCAGGACAGATTAATCCCGGCGGAGAGA 1060
172 lalatrparargluprotprilleshisalarproglnglycysgly 188
1061 GGGTGTAGGGCGAGAG.....ATCACCGTCACCGGCGATATGCGAG 1104
189 AsnSerSerArgThrIleThrCysAspThr 198
1105 AACACGCGACAG.....ATCGCGCATACG 1128

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-096-080-27

```



```

; Sequence 26535, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 202-02-21
; SEQ ID NO 26535
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Synecocystis sp.
US-60-360-039-26535

```

```

alignment_scores:
  Quality: 77.00      Length: 201
  Ratio: 0.885      Gaps: 7
  Percent Similarity: 43.284      Percent Identity: 21.891

```

alignment_block:

US-09-528-682-1 x US-60-360-039-26535 ..

Align seg 1/1 to: US-60-360-039-26535 from: 1 to: 2163

```

15 GluIleLysArgSerGlyLeuMetPro.....ArgLysH 27
   ||||| ||| ||||| |||||
1624 GAATTTATTTCTCCCGGATTTGATCCCGACAGATTGGCGCATCA 1673
27 sAaNGlUtyrPhaSPaRgSlYthrGlnMetAsnIleAsnLeuTyrAspH 44
   - - - - - ||||| |||||
1674 CATGATTAATTTG.....TGGAAACCGTAACGCCATTTGTCG 1708
44 tSaLaArgGlyThrGlnThrGlyPheValArgTyrAspAspGlyTyrVal 60
   ||||| ||||| |||||
1709 GTATTTGGGGGAATTCACCGGTTTACC..... 1737
61 SerThrSerLeuSerLeuArgSerAlaHisIleuAlaGlyInSerIleLe 77
   ||||| ||||| |||||
1738 .....CTCCAAACCCAAAGGAAGAACT 1757
77 uSerGlyTyrSerThrTyrTyrTyrTyrValIleAlaThrAlaProAsnM 94
   ||||| ||||| ||||| |||||
1758 GAGCGCTATTAATAATTAGCTATTGTAAGTCGCGCATCGCCCTTACC 1807
94 eTpHeAsnValAsnAspValIleuGlyValTyrSerProHisProTyrGlu 110
   ||||| ||||| |||||
1808 TAGAGAAATTCGCAACTGTG.....CATTAAGGGGAA 1842
111 GlnGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTr 127
   ||||| ||||| |||||
1843 CAACACCTTAATACATC..... 1860
127 pTyArgValAsnPhaGlyValIleAspGluArgLeu.....HisArg 142
   ||||| ||||| |||||
1861 .....ATTTCCTTTGGTCAATTTTGGATTTCTGCGCATCGCACTT 1903
142 sNaArgLyuTyArgAspArgTyrTyrTyrArgAsnLeuAsnIleAlaProAl 158
   ||||| ||||| |||||
1904 CCAAGGCTATGCTGCTGCTGTTGATGACCAAGTGGAAATATTCCTCTG 1953
159 GluAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrPar 175
   ||||| ||||| |||||
1954 GAGCAGCTTTTCAACCGCGAGATCGGAGCGGACGACAGATATGATGCG 2003
175 gGluGluProTyrPheHisIleAlaProGlnGlyCysGlyAsnSerSera 192
   ||||| ||||| |||||
2004 G.....GGTAACACCTTT 2017

```

```

192 rGthrIleThrGlyAspThrCysAsnGluGlyThrGlnAsnLeuSerThr 208
   ||||| ||||| |||||
2018 CGTCTGCTGCGTACACGTCACACATGAGAACTTTCTATATCAGGGAG 2067
209 Ile 209
   |||
2068 ATC 2070

```

seq_name: /cgn2_6/plodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-1307

seq_documentation_block:

```

; Sequence 1307, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1307
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1307

```

```

alignment_scores:
  Quality: 77.00      Length: 187
  Ratio: 0.802      Gaps: 11
  Percent Similarity: 51.337      Percent Identity: 23.529

```

alignment_block:

US-09-528-682-1 x US-10-104-047-1307 ..

Align seg 1/1 to: US-10-104-047-1307 from: 1 to: 2405

```

9 AspSerArgProProAspGluIleLysArgSerGlyLeuMetProThr 25
   ||||| ||||| ||||| |||||
1465 GACACGTGCTCCCGCTGACATCTACCCCAAGATCTGACCTCCACCGG 1514
25 gGlnHis...AsnGlyTyrPheAspArgGlyThrGlnMetAsnIleAsnL 41
   ||||| ||||| ||||| |||||
1515 TCACCATGGCAACCTTACACATGACCGCTGCGCTGCTGCTGCC 1563
41 euTyArgAspHis..AlaArgGlyThrGlnThrGlyPheValArgTyrAsp 57
   ||||| ||||| ||||| |||||
1564 ..TATGATCATCGCGCGTGGCAAT.....GACCA 1590
57 pGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisIleuAlaGly 74
   ||||| ||||| ||||| |||||
1591 GTGGCCATCAGCTCCAAATTTGACACCGGGAG.....GATA 1628
74 InsrIleLeuSerGlyTyrSerThrTyrTyrTyrValIleAlaThr 90
   ||||| ||||| ||||| |||||
1629 TTGCTGTGATCCGAGACTATGGAACTCTCTGTGAGATGTCGCTG 1678
91 AlaProAsnMetPheAsnValAsnAspValIleuGlyValTyrSerProH 107
   ||||| ||||| ||||| |||||
1679 GTCCCT.....GATGGCATCGGGCTTCTTACACGAGTA 1713
107 sProTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnI 124
   ||||| ||||| ||||| |||||
1714 CCAGTACATGAG.....AGCACCG 1733
124 leTyArgLysTyrTyrArgValAsnPhaGlyValIleAspGluArgLeuH 140
   ||||| ||||| ||||| |||||
1734 TGGCTCTCTGTTGAGCAG.....GGGATCTT...GAGAACTCCAG 1774
141 ArgAsnArgGlyTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaPr 157

```


[illegible]

OM of: US-09-528-682-1 to: GenEmbl:* out_format : pfs
Date: Jun 18, 2002 10:49 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09528682/runat_18062002.082606.8807/app_query.fasta.1.299
-DB=genembl -QMT=fastap -SUFFIX=olip2n.rge -GAPOP=4.500
-GAPXPT=0.050 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPXPT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-DELOP=6.000 -DELEXT=7.000 -XGAPOP=60.000 -XGAPEXT=60.000
-TRANS=human40.cdi -LIST=1000 -DOCALLIG=200 -THR.SCORE=quality
-THR.MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM=ext
-HEASIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09528682.EGNC1.1.6842 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPRX -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-1
Query Length: 240
Database: GenEmbl:*
Database sequences: 1797656
Database Length: 1673333701
Search time (sec): 1839.660000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd Orig	ZScore	EScore Len	Documentation
gb_da:AF242417	+	240.00	4660.07	3.3e-251 777 AF242417 Escherichia coli heat
gb_da:P307LTRA	+	240.00	4660.07	3.3e-251 777 M35581 Plasmid P307 (from E. coli
gb_da:AE04913	+	240.00	4660.07	3.3e-251 777 A04913 E. coli LTR gene for labi
gb_da:ENSTRTA	+	240.00	4659.29	3.7e-251 866 M15362 Plasmid ENT p307 (enter
gb_da:ENRTTPA	+	240.00	4659.03	3.8e-251 898 M15361 Plasmid ENT-R pCG86 (ent
gb_da:EB01421	+	208.00	4332.65	2.9e-216 1148 I A03421 DNA encoding Lth of ent
gb_da:EB011677	+	191.00	3699.22	1.1e-197 1434 I AB011677 Escherichia coli gene
gb_da:ECOTXOA	+	184.00	3563.42	4.0e-190 934 I K01995 E. coli toxa gene encodi
gb_da:SE0731	+	184.00	3563.42	4.0e-190 1275 I SE0731 heat labile enterotoxin
gb_pat:AE04915	+	179.00	3469.35	7.0e-185 777 I A04915 Artificial sequence LTR
gb_da:EMRTA	+	111.00	2142.52	5.6e-111 723 I M57244 Plasmid BMD 299 (from E.
gb_pat:EE0613	+	97.00	1868.72	1.0e-95 777 I E0613 DNA encoding A component
gb_da:AE16419	+	86.00	1654.64	8.4e-84 711 I A16419 A subunit of a labile to
gb_pat:AE18597	+	86.00	1653.22	1.0e-83 868 I V00275 E. coli gene elta encodi
gb_da:ECRTA	+	34.00	640.88	2.5e-37 595 I X8396 E. coli LTR7 gene for hea
gb_pat:AE06082	+	30.00	560.77	7.2e-23 791 I A06082 PEK620 HindII-Fragment.
gb_pat:AE06083	+	30.00	560.77	7.2e-23 791 I A06083 PEK620 HindII-Fragment.
gb_pat:AE0701	+	30.00	560.73	7.2e-23 795 I A0701 Synthetic DNA sequence f
gb_pat:AE0702	+	30.00	560.73	7.2e-23 795 I A0702 Synthetic DNA sequence f
gb_da:AF45254	+	27.00	502.33	1.3e-19 777 I AF45254 Vibrio cholerae strain
gb_da:AF463401	+	27.00	502.33	1.3e-19 777 I AF463401 Vibrio cholerae strain
gb_da:AF63401	+	27.00	502.33	1.3e-19 777 I AF63401 Vibrio cholerae strain
gb_da:VIBCTXA1	+	27.00	498.28	2.2e-19 1372 I B30052 Vibrio cholerae ctxa ge
gb_da:VIBCTXA1	+	26.00	483.33	1.5e-18 723 I A16422 A subunit of cholera tox
gb_pat:AE18598	+	26.00	483.33	1.5e-18 723 I A18598 Sequence 7 from patent
gb_da:VIBCTXA	+	26.00	482.81	1.6e-18 777 I K03679 Vibrio cholerae enteroc
gb_pat:EE0132	+	26.00	480.03	2.2e-18 1148 I E0132 DNA coding of cholera t
gb_da:VICTOXA21	+	26.00	478.81	2.6e-18 1362 I X8786 V. cholerae genes for to
gb_da:VIBCTXA5	+	26.00	478.77	2.6e-18 1369 I D30053 Vibrio cholerae ctxa ge
gb_da:VICTOXA56	+	26.00	478.73	2.7e-18 1397 I X8785 V. cholerae genes for to
gb_da:VCCXTX	+	26.00	475.99	3.8e-18 2020 I X00171 Vibrio cholera toxin (C
gb_pat:AE05312	+	26.00	475.99	3.8e-18 2020 I E03512 Cholera toxin gene. 9/1
gb_da:AF175708	+	26.00	472.11	6.2e-18 3480 I AF175708 Vibrio cholerae acces
gb_da:AF14359	+	26.00	470.64	7.5e-18 4275 I AF14359 Vibrio phage CTX Ace
gb_da:AE004224	+	26.00	464.13	1.7e-17 10646 I AE004224 Vibrio cholerae chrc
gb_da:ECCELTR72	+	24.00	454.83	5.7e-17 165 I K0043 E. coli heat labile enter
gb_da:ECCELTR	+	24.00	443.80	2.3e-16 774 I J01646 E. coli heat labile enter

gb_da:ECOPHLAGS	+	22.00	403.68	4.0e-14	900 I M61015 Escherichia coli heat
gb_da:VIBCTA1	+	13.00	237.01	7.7e-05	255 I J01846 Vibrio cholerae chole
gb_da:VIBCTA2	+	12.00	217.60	0.0009	251 I J01847 Vibrio cholerae chole
gb_da:VIBCTOX	+	11.00	191.81	0.0018	512 I M1770 Vibrio cholerae toxa
gb_da:ECOCERTXBL	+	11.00	186.55	0.0254	604 I M1873 E. coli (from p19) hea
gb_da:ECOCERTXHL	+	11.00	186.55	0.0499	1262 I M28523 Escherichia coli (st
gb_pat:AR019034	+	10.00	191.32	0.0271	42 I AR019034 Sequence 14 from pat
gb_pat:AE69697	+	10.00	161.54	1.23	2734 I AE69697 Sequence 15 from pat
gb_pat:BD009884	+	10.00	161.54	1.23	2734 I BD009884 Recombinant toxin
gb_htg:AC068893	+	10.00	132.72	49.76	154611 I AC068893 Homo sapiens chr
gb_htg:AC021033	+	10.00	132.18	53.36	166872 I AC021033 Homo sapiens chr
gb_pat:AC073367	+	10.00	132.09	53.96	168919 I AC073367 Homo sapiens chr
gb_htg:AC021030	+	10.00	131.52	58.07	183028 I AC021030 Homo sapiens chr
gb_htg:AC021031	+	10.00	131.38	59.10	186581 I AC021031 Homo sapiens chr
gb_htg:AC021041	+	10.00	131.00	62.06	196822 I AC021041 Homo sapiens chr
gb_htg:AC087184	+	10.00	130.66	64.82	206400 I AC087184 Mus musculus chr
gb_pat:AE03500	+	9.00	174.95	0.2211	27 I E03500 primer for amplifying
gb_pat:AE16408	+	9.00	174.20	0.2435	30 I AE16408 primer. 9/1994
gb_pat:AE03497	+	9.00	173.10	0.2804	35 I E03497 primer for amplifying
gb_da:VCCCTXAB2	+	9.00	150.24	5.26	861 I X76391 V. cholerae (0133) ctx
gb_da:VCCCTXAB	+	9.00	149.63	5.69	938 I X76390 V. cholerae (0133) ctx
gb_da:AE009014	+	9.00	132.54	50.94	10291 I AE009014 Agrobacterium tum
gb_htg:AC108816	+	9.00	130.45	66.58	13788 I AE007980 Agrobacterium tum
gb_htg:AC108816	+	9.00	120.79	229.72	53355 I AC108816 Mus musculus cton
gb_htg:AC092373	+	9.00	114.52	513.44	128466 I AC092373 Homo sapiens chr
gb_htg:AC061557	+	9.00	113.56	580.61	146936 I AC106157 Rattus norvegicu
gb_htg:AC073176	+	9.00	112.30	682.95	175454 I AC073176 Homo sapiens chr
gb_htg:AC023639	+	9.00	111.85	723.51	186868 I AC023639 Homo sapiens chr
gb_htg:AC106284	+	9.00	111.35	771.04	200319 I AC106284 Rattus norvegicu
gb_htg:AC073368	+	9.00	109.51	976.16	259210 I AC073368 Homo sapiens chr
gb_pat:AE03631	+	8.00	156.27	2.43	24 I E03631 DNA primer for detecti
gb_pat:AR118626	+	8.00	155.17	2.79	28 I AR118626 Sequence 36 from pat
gb_pat:AE004452	+	8.00	139.42	3.16	32 I AE004452 D. ancesis internal tra
gb_pl:DA28444	+	8.00	139.49	20.88	252 I DA28444 human STRS WI-3130, se
gb_sts:G37193	+	8.00	138.44	23.89	292 I G04193 human STRS WI-3130, se
gb_sts:G37193	+	8.00	136.32	31.36	393 I G37193 SHGC-57937 Human Homo
em_hlgo:AC029667	+	8.00	135.53	34.70	439 I AC029667 Giardia intestinali
gb_pl:BLVACPIIT	+	8.00	134.32	40.52	520 I M24426 Barley acyl carrier p
gb_pat:AF282562	+	8.00	133.70	43.86	567 I AF282562 Homo sapiens clone
gb_sts:G59212	+	8.00	133.28	46.33	602 I G59212 SHGC-81474 Human Homo
gb_sts:G59212	+	8.00	132.52	51.03	669 I G59212 SHGC-110146 Human Hom
em_hlgo:AC032116	+	8.00	132.09	53.96	711 I AC032116 Giardia intestinali
gb_pat:HSA337570	+	8.00	131.79	56.04	741 I AJ337370 Homo sapiens genom
em_hlgo:AC074680	+	8.00	131.66	57.00	755 I AC074680 Giardia intestinali
em_hlgo:AC051599	+	8.00	130.75	64.08	858 I AC051599 Giardia intestinali
em_hlgo:AC028724	+	8.00	130.72	64.29	861 I AC028724 Giardia intestinali
em_hlgo:AC067411	+	8.00	129.56	65.65	881 I AC067411 Giardia intestinali
em_hlgo:AC065255	+	8.00	129.96	70.89	958 I AC065255 Giardia intestinali
em_hlgo:AC045117	+	8.00	129.70	73.25	993 I AC045117 Giardia intestinali
em_hlgo:AC030359	+	8.00	129.68	73.46	996 I AC030359 Giardia intestinali
em_hlgo:AC057682	+	8.00	128.97	80.45	1100 I AC057682 Giardia intestinali
gb_pat:AK026292	+	8.00	127.57	96.04	1335 I AK026292 Homo sapiens cDNA:
gb_pat:BC017652	+	8.00	127.56	96.44	1341 I BC017652 Homo sapiens, hypo
gb_pat:BC017652	+	8.00	125.09	116.53	1649 I AC035578 Amblyomma american
gb_hlgo:AC021349	+	8.00	125.16	131.20	1877 I BC021349 Mus musculus, cton
gb_pat:BC016357	+	8.00	123.88	154.56	2245 I BC016357 Homo sapiens, cton
gb_pat:AE012153	+	8.00	123.82	155.81	2265 I AE012153 Sequence 1273 from
gb_vl:AHVCPV3A	+	8.00	122.34	183.32	2766 I M94312 African horseshoeknes
gb_vl:AHVCPV3A	+	8.00	122.33	188.51	2769 I D26572 African horseshoeknes
gb_vl:AHVCPV3A	+	8.00	122.33	188.69	2792 I AF021236 African horse sick
gb_vl:AHVCPV3A	+	8.00	122.33	188.69	2792 I M94681 African horseshoeknes
gb_vl:AHVCPV3A	+	8.00	121.78	202.38	3014 I BC004766 Mus musculus, slmt
gb_vl:AHVCPV3A	+	8.00	121.62	206.62	3063 I AC069512 Drosophila melanog
gb_pat:AE03850	+	8.00	119.88	258.38	3936 I E30850 Cytochrome bd type q
gb_pat:AE03850	+	8.00	119.88	258.38	3936 I E30851 Cytochrome bd type q
gb_da:AB030366	+	8.00	119.87	258.86	3939 I AB030366 Corynebacterium gl
gb_da:HSU01036	+	8.00	119.74	262.85	4011 I HSU01036 Helicobacter sternalis
gb_vl:RVU07141	+	8.00	118.78	297.30	4580 I U17141 Powloox virus ORF1 a
gb_pl:BLVACPIIT	+	8.00	118.42	311.62	4830 I M8754 Hordeum vulgare acyl
gb_pat:AE0097461	+	8.00	115.66	443.94	7110 I AE0097461 Sequence 26 from p
gb_in:DMSALTR	+	8.00	115.04	480.38	7750 I Y07653 D. melanogaster spalt
gb_da:AE006837	+	8.00	112.33	680.32	11335 I AE006837 Sulfolobus solfat

gb_ba:AE004158	+	8.00	112.30	682.68	11378	AE004158	Vibrio cholerae chcd	gb_htg:LMFLCHR32_22	+	8.00	96.11	5.4e+03	110000	Continuation (23 of 36) o
gb_ba:AE006663	+	8.00	112.96	685.92	11437	AE006663	Sulfolobus solitarius	gb_htg:LMFLCHR36_24	+	8.00	96.11	5.4e+03	110000	Continuation (25 of 36) o
gb_in:CEH19N07	+	8.00	111.94	714.90	11966	292835	Caenorhabditis elegans	gb_pl:AC006258	+	8.00	96.08	5.5e+03	110469	AC006258 Arabidopsis thal
gb_ba:AE009278	-	8.00	111.59	747.59	12555	AE009278	Agrobacterium tumefaciens	gb_htg:AP001801	+	8.00	96.03	5.5e+03	111187	AP001801 Homo sapiens chr
gb_htg:AC014206	-	8.00	111.49	757.28	12743	AC014206	Drosophila melanogaster	gb_pr:AC026441	+	8.00	95.95	5.6e+03	112532	AC026441 Homo sapiens chr
gb_ba:AE001969	+	8.00	111.35	771.62	13007	AE001969	Deinococcus radiodurans	gb_pr:AC010584	+	8.00	95.90	5.6e+03	113216	AC010584 Homo sapiens chr
gb_ba:AE008334	+	8.00	110.45	861.18	14665	AE008334	Agrobacterium tumefaciens	gb_pr:AL133545	+	8.00	95.86	5.6e+03	113988	AL133545 Homo sapiens chr
gb_ba:AE004755	-	8.00	110.44	866.23	14759	AE004755	Pseudomonas aeruginosa	gb_htg:AC103101	+	8.00	95.73	5.7e+03	116089	AC103101 Rattus norvegicus
gb_htg:AC109866	+	8.00	109.78	979.73	16884	AC109866	Rattus norvegicus cl	gb_pr:AL15780	+	8.00	95.65	5.8e+03	116878	AL15780 Homo sapiens chr
gb_htg:AC109707	+	8.00	107.45	1.3e+03	21584	AC109707	Rattus norvegicus cl	gb_htg:AC008457	+	8.00	95.65	5.8e+03	117366	AC008457 Homo sapiens chr
gb_pat:AX076880	+	8.00	107.32	1.3e+03	22870	AX076880	Sequence 1 from Patk	gb_pl:AC007534	+	8.00	95.58	5.9e+03	118444	AC007534 Rattus norvegicus
gb_pr:AC004462	+	8.00	105.04	1.7e+03	31491	AC004462	Sequence 1 from Patk	gb_htg:AC009759	+	8.00	95.55	5.9e+03	118968	AC009759 Rattus norvegicus
gb_pr:HS366D1	+	8.00	104.08	2.0e+03	36021	HS366D1	Homo sapiens Chromos	gb_htg:AF198096	+	8.00	95.45	5.9e+03	120731	AF198096 Homo sapiens chr
gb_htg:AC019927	+	8.00	103.51	2.0e+03	36879	AC019927	Drosophila melanogaster	gb_htg:AP001078	+	8.00	95.40	6.0e+03	121448	AP001078 Homo sapiens chr
gb_pr:AC005061	-	8.00	103.62	2.1e+03	38413	AC005061	Homo sapiens Chromos	gb_pr:AC016648	+	8.00	95.36	6.0e+03	122218	AC016648 Homo sapiens chr
gb_pl:AC005299	-	8.00	103.23	2.2e+03	40554	AC005299	emericella nidulans	gb_pr:AC004904	+	8.00	95.14	6.2e+03	126041	AC004904 Homo sapiens chr
gb_htg:AC096021	+	8.00	103.03	2.2e+03	41788	AC096021	Rattus norvegicus cl	gb_htg:AC099140	+	8.00	95.10	6.3e+03	128511	AC099140 Rattus norvegicus
gb_htg:AC110114	+	8.00	102.55	2.4e+03	44615	AC110114	Rattus norvegicus cl	gb_htg:AC096871	+	8.00	94.88	6.4e+03	130692	AC096871 Rattus norvegicus
gb_pl:R22M8	-	8.00	102.13	2.5e+03	47295	AC020622	Sequence of BAC F22M	gb_pr:AC079772	+	8.00	94.88	6.4e+03	130743	AC079772 Homo sapiens chr
gb_htg:AC006182	-	8.00	101.79	2.6e+03	49636	AC006182	Homo sapiens Chromos	gb_pr:AC008971	+	8.00	94.82	6.4e+03	131856	AC008971 Homo sapiens chr
gb_pl:AV013245	-	8.00	101.79	2.6e+03	50000	AV013245	Oryza sativa Chromos	gb_htg:AF205589	+	8.00	94.67	6.5e+03	134570	AF205589 Homo sapiens chr
gb_pr:AL356515	-	8.00	100.81	3.0e+03	56951	AL356515	Homo sapiens Chromos	gb_htg:AL356692	+	8.00	94.66	6.5e+03	134787	AL356692 Homo sapiens chr
gb_htg:AC017426	-	8.00	100.47	3.1e+03	59704	AC017426	Drosophila melanogaster	gb_htg:AC103239	+	8.00	94.50	6.7e+03	137897	AC103239 Rattus norvegicus
gb_pr:AC000386	-	8.00	100.43	3.1e+03	60023	AC000386	Homo Chromosome 15	gb_pr:AL157764	+	8.00	94.48	6.7e+03	138251	AL157764 Homo sapiens chr
gb_pr:AC000260	+	8.00	100.04	3.3e+03	63402	AP000260	Homo sapiens genom	gb_pr:AC004908	+	8.00	94.48	6.7e+03	138251	AC004908 Homo sapiens chr
gb_pr:HS1173K17	-	8.00	99.89	3.4e+03	64810	AL050315	Homo sapiens genom	gb_pr:AC009226	+	8.00	94.45	6.7e+03	138778	AC009226 Homo sapiens chr
gb_htg:AC090157	+	8.00	99.84	3.4e+03	65259	AC090157	Homo sapiens Chromos	gb_pl:AL157821	+	8.00	94.45	6.7e+03	138781	AL157821 Homo sapiens chr
gb_htg:AC090157	+	8.00	99.84	3.4e+03	65259	AC090157	Homo sapiens Chromos	gb_pr:ATAC009177	+	8.00	94.39	6.8e+03	140064	ATAC009177 Arabidopsis thal
gb_htg:AC100935	-	8.00	99.71	3.4e+03	66445	AC100935	Homo sapiens Chromos	gb_pr:AC001253	+	8.00	94.38	6.8e+03	140249	AC001253 Homo sapiens chr
gb_htg:AC110274	+	8.00	99.49	3.5e+03	68482	AC110274	Homo sapiens Chromos	gb_htg:AC013376	+	8.00	94.37	6.8e+03	140290	AC013376 Homo sapiens chr
gb_pr:HSU70984	+	8.00	99.46	3.5e+03	68548	U70984	Human X22 cosmid	gb_pr:AL356107	+	8.00	94.36	6.8e+03	140857	AL356107 Homo sapiens chr
gb_htg:AC090318	-	8.00	99.46	3.5e+03	68733	AC090318	Homo sapiens Chromos	gb_pr:AP000036	+	8.00	94.15	7.0e+03	144857	AP000036 Homo sapiens chr
gb_htg:AC107379	-	8.00	99.38	3.6e+03	69537	AC107379	Homo sapiens Chromos	gb_htg:AC069591	+	8.00	94.12	7.0e+03	145458	AC069591 Homo sapiens chr
gb_htg:AC027823	-	8.00	99.26	3.6e+03	70720	AC027823	Homo sapiens Chromos	gb_pr:AC008766	+	8.00	94.08	7.1e+03	147530	AC008766 Homo sapiens chr
gb_htg:AC101550	+	8.00	99.18	3.7e+03	71588	AC101550	Homo sapiens Chromos	gb_htg:AC057542	+	8.00	93.92	7.2e+03	148172	AC057542 Fells calus chr
gb_htg:AC025914	+	8.00	99.15	3.7e+03	72450	AC025914	Mus musculus Chromos	gb_htg:AF166490	+	8.00	93.97	7.2e+03	148355	AF166490 Homo sapiens chr
gb_pl:AC074228	+	8.00	99.05	3.7e+03	72832	AC074228	Arabidopsis thaliana	gb_htg:AF1627305	+	8.00	93.93	7.2e+03	149354	AF1627305 Drosophila melan
gb_pr:AL136159	+	8.00	98.94	3.8e+03	74012	AL136159	Homo sapiens Chromos	gb_htg:AP004193	+	8.00	93.88	7.2e+03	150275	AP004193 Oryza sativa chr
gb_pr:AP001264	+	8.00	98.89	3.8e+03	74550	AP001264	Homo sapiens genom	gb_htg:AL356215	+	8.00	93.87	7.3e+03	150533	AL356215 Homo sapiens chr
gb_htg:AC037461	+	8.00	98.59	4.0e+03	77654	AC037461	Homo sapiens Chromos	gb_htg:AC018889	+	8.00	93.87	7.3e+03	150629	AC018889 Bos taurus clone
gb_pr:AL590611	-	8.00	98.57	4.0e+03	77888	AL590611	Homo sapiens Chromos	gb_pr:AC007719	+	8.00	93.86	7.3e+03	150831	AC007719 Homo sapiens chr
gb_pr:AL139826	-	8.00	98.35	4.1e+03	80419	AL139826	Homo sapiens Chromos	gb_pr:AL354740	+	8.00	93.81	7.3e+03	151828	AL354740 Homo sapiens chr
gb_pr:AC084807	+	8.00	98.23	4.1e+03	81704	AC084807	Arabidopsis thaliana	gb_pr:AC0359074	+	8.00	93.79	7.3e+03	152161	AC0359074 Homo sapiens chr
gb_pr:AC334726	+	8.00	98.18	4.2e+03	82316	AL334726	Homo sapiens Chromos	gb_htg:AC103906	+	8.00	93.76	7.3e+03	152230	AC103906 Homo sapiens chr
gb_pl:AC007197	-	8.00	98.09	4.2e+03	83363	AC007197	Arabidopsis thaliana	gb_htg:AC011988	+	8.00	93.74	7.4e+03	153386	AC011988 Homo sapiens chr
gb_htg:AC017221	+	8.00	98.05	4.2e+03	83824	AC017221	Drosophila melanogaster	gb_htg:AC087656	+	8.00	93.71	7.4e+03	154067	AC087656 Homo sapiens chr
gb_pr:AP001255	+	8.00	98.04	4.3e+03	83921	AP001255	Homo sapiens genom	gb_htg:AC026033	+	8.00	93.68	7.4e+03	154323	AC026033 Homo sapiens chr
gb_htg:AL645826	+	8.00	97.62	4.5e+03	89064	AL645826	Danio rerio Chromos	gb_pr:AL355871	+	8.00	93.68	7.4e+03	154588	AL355871 Homo sapiens chr
gb_htg:AC105557	+	8.00	97.55	4.5e+03	89964	AC105557	Rattus norvegicus cl	gb_pr:AC004846	+	8.00	93.68	7.4e+03	154616	AC004846 Homo sapiens chr
gb_htg:AC109441	+	8.00	97.54	4.5e+03	89990	AC109441	Homo sapiens Chromos	gb_htg:AC026059	+	8.00	93.67	7.4e+03	154630	AC026059 Homo sapiens chr
gb_pr:AC008954	-	8.00	97.53	4.5e+03	90111	AC008954	Homo sapiens Chromos	gb_htg:AC011859	+	8.00	93.67	7.4e+03	154809	AC011859 Homo sapiens chr
gb_pr:AF225900	+	8.00	97.18	4.7e+03	94645	AF225900	Homo sapiens PAC cld	gb_pr:AC092782	+	8.00	93.65	7.5e+03	155355	AC092782 Homo sapiens chr
gb_pr:HSB273H14	+	8.00	97.11	4.8e+03	95556	AL080272	Homo sapiens Chromos	gb_htg:AC015953	+	8.00	93.62	7.5e+03	155904	AC015953 Homo sapiens chr
gb_htg:AC016004	+	8.00	97.06	4.8e+03	96251	AC016004	Homo sapiens clone	gb_in:AC023748	+	8.00	93.61	7.5e+03	156509	AC023748 Drosophila melan
gb_pr:HS633020	+	8.00	97.05	4.8e+03	96460	AL022804	Homo sapiens clone	gb_pr:AC023748	+	8.00	93.61	7.5e+03	156509	AC023748 Drosophila melan
gb_htg:AC106033	-	8.00	97.03	4.8e+03	96732	AC106033	Homo sapiens Chromos	gb_pr:AC021914	+	8.00	93.58	7.5e+03	156829	AC021914 Homo sapiens chr
gb_htg:AC105539	+	8.00	96.92	4.9e+03	97307	AC105539	Rattus norvegicus cl	gb_pr:AL450333	+	8.00	93.57	7.5e+03	156964	AL450333 Homo sapiens chr
gb_htg:AC096236	-	8.00	96.99	4.9e+03	98211	AC096236	Rattus norvegicus cl	gb_htg:AL670729	+	8.00	93.57	7.5e+03	157033	AL670729 Homo sapiens chr
gb_htg:AP000685	-	8.00	96.91	4.9e+03	98305	AP000685	Homo sapiens Chromos	gb_htg:AC087109	+	8.00	93.57	7.5e+03	157121	AC087109 Homo sapiens chr
gb_htg:AL592068	+	8.00	96.90	4.9e+03	98454	AL592068	Homo sapiens Chromos	gb_pr:AC092670	+	8.00	93.53	7.6e+03	157951	AC092670 Homo sapiens chr
gb_pr:AP000099	+	8.00	96.79	5.0e+03	100000	AP000099	Homo sapiens genom	gb_pr:AC011155	+	8.00	93.51	7.6e+03	158371	AC011155 Homo sapiens chr
gb_pr:AP000175	+	8.00	96.79	5.0e+03	100000	AP000175	Homo sapiens genom	gb_htg:AC091404	+	8.00	93.49	7.6e+03	158755	AC091404 Sus scrofa clon
gb_pr:AP003087	+	8.00	96.76	5.0e+03	100403	AP003087	Homo sapiens genom	gb_pr:AC006013	+	8.00	93.48	7.6e+03	159002	AC006013 Homo sapiens chr
gb_pr:AL390066	+	8.00	96.54	5.2e+03	103611	AL390066	Homo sapiens genom	gb_htg:AC068293	+	8.00	93.47	7.6e+03	159339	AC068293 Homo sapiens chr
gb_htg:AC014357	+	8.00	96.53	5.2e+03	103746	AC014357	Drosophila melanogaster	gb_pr:AC074008	+	8.00	93.46	7.6e+03	159439	AC074008 Homo sapiens chr
gb_htg:AC094679	+	8.00	96.27	5.3e+03	107637	AC094679	Rattus norvegicus	gb_pr:AC022919	+	8.00	93.45	7.7e+03	159749	AC022919 Homo sapiens chr
gb_pr:AC094679	+	8.00	96.21	5.4e+03	108400	L77570	Homo sapiens Digeorg	gb_htg:AC099783	+	8.00	93.44	7.7e+03	159870	AC099783 Homo sapiens chr
gb_pr:AC026736	+	8.00	96.20	5.4e+03	108661	AC026736	Homo sapiens Chromos	gb_pr:AC009082	+	8.00	93.42	7.7e+03	160277	AC009082 Homo sapiens chr
gb_pr:AC083949	-	8.00	96.16	5.4e+03	110926	AC083949	Homo sapiens Chromos	gb_htg:AC096203	+	8.00	93.42	7.7e+03	160424	AC096203 Rattus norvegicus
gb_htg:AC096666_1	-	8.00	96.11	5.4e+03	110000	Continuation (2 of 4) of AC0		gb_htg:AC098166	+	8.00	93.42	7.7e+03	160424	AC098166 Rattus norvegicus
gb_htg:AL627224_1	+	8.00	96.11	5.4e+03	110000	Continuation (2 of 5) of AL6		gb_htg:AC068372	+	8.00	93.42	7.7e+03	160426	AC068372 Homo sapiens chr

gb_hhg:AC021933	+	8.00	93.40	7.7e+03	160864	AC021933	Homo sapiens chrom	gb_pr:AC092724	+	8.00	92.73	8.4e+03	176521	AC092724	Homo sapiens chr
gb_pr:AC010138	-	8.00	93.39	7.7e+03	160970	AC010138	Homo sapiens chrom	gb_in:AC008029	+	8.00	92.72	8.4e+03	176591	AC008029	Drosophila melan
gb_hhg:AC087464	-	8.00	93.38	7.7e+03	161262	AC087464	Homo sapiens chrom	gb_pr:AF199033	+	8.00	92.71	8.4e+03	177026	AF199033	Homo sapiens chr
gb_pr:HS169M13	+	8.00	93.38	7.7e+03	161273	AL035427	Human DNA sequence	gb_hhg:AC096334	-	8.00	92.70	8.4e+03	177276	AC096334	Rattus norvegicus
gb_hhg:AC024229	-	8.00	93.37	7.7e+03	161540	AC024229	Homo sapiens chrom	gb_hhg:AL626782	-	8.00	92.70	8.4e+03	177475	AL626782	Mus musculus chr
gb_hhg:AC0106148	-	8.00	93.37	7.7e+03	161542	AC0106148	Rattus norvegicus c	gb_hhg:AC098062	-	8.00	92.70	8.4e+03	177509	AC098062	Rattus norvegicus
gb_pr:AL192045	+	8.00	93.35	7.8e+03	161933	AL192045	Human DNA sequence	gb_hhg:AC026906	-	8.00	92.68	8.5e+03	177904	AC026906	Homo sapiens chr
gb_hhg:AC002986	+	8.00	93.34	7.8e+03	162066	AP002986	Homo sapiens chrom	gb_hhg:AC065936	-	8.00	92.67	8.5e+03	178073	AC065936	Rattus norvegicus
gb_hhg:AC019271	+	8.00	93.34	7.8e+03	162167	AC019271	Homo sapiens clone	gb_pr:AL930723	+	8.00	92.67	8.5e+03	178223	AL930723	Human DNA sequen
gb_hhg:AC009367	+	8.00	93.31	7.8e+03	162851	AC009367	Homo sapiens chrom	gb_hhg:AC009420	-	8.00	92.67	8.5e+03	178229	AC009420	Homo sapiens clo
gb_hhg:AC0091925	+	8.00	93.31	7.8e+03	162914	AC0091925	Homo sapiens chrom	gb_hhg:AC0011968	-	8.00	92.66	8.5e+03	178313	AC0011968	Homo sapiens chr
gb_hhg:AC078804	+	8.00	93.29	7.8e+03	163243	AC078804	Homo sapiens chrom	gb_hhg:AL645822	-	8.00	92.66	8.5e+03	178313	AL645822	Danio rerio clon
gb_pr:AC093868	+	8.00	93.29	7.8e+03	163334	AC093868	Homo sapiens chrom	gb_hhg:AC016383	+	8.00	92.65	8.5e+03	178680	AC016383	Homo sapiens chr
gb_hhg:AC011332	+	8.00	93.26	7.8e+03	163059	AC011332	Homo sapiens clone	gb_pr:AC079316	+	8.00	92.64	8.5e+03	178905	AC079316	Homo sapiens 12q
gb_pr:AC068614	+	8.00	93.26	7.9e+03	164059	AC068614	Rattus norvegicus c	gb_hhg:AP001934	+	8.00	92.62	8.5e+03	179281	AP001934	Homo sapiens chr
gb_hhg:AC026657	+	8.00	93.23	7.9e+03	164706	AC026657	Homo sapiens BAC cl	gb_hhg:AC0090107	+	8.00	92.62	8.5e+03	179470	AC0090107	Homo sapiens chr
gb_pr:CNS010UN	+	8.00	93.23	7.9e+03	164766	AC026657	Homo sapiens chrom	gb_hhg:AC0022177	+	8.00	92.59	8.6e+03	180076	AC022177	Homo sapiens clo
gb_hhg:AC009289	-	8.00	93.22	7.9e+03	164935	AL133329	Human chromosome 14	gb_hhg:AC104820	+	8.00	92.58	8.6e+03	180393	AC104820	Homo sapiens chr
gb_pr:CNS01DRF	-	8.00	93.22	7.9e+03	164958	AC092729	Human chromosome 14	gb_pr:CNS01DRF	-	8.00	92.57	8.6e+03	180523	AC092729	Human chromosome
gb_hhg:AC0092729	-	8.00	93.21	7.9e+03	165049	AL117190	Human chromosome 14	gb_hhg:AC022080	-	8.00	92.57	8.6e+03	180676	AC022080	Homo sapiens chr
gb_pr:CNS01DRF	+	8.00	93.21	7.9e+03	165159	AL133454	Human chromosome 14	gb_hhg:AC104299	+	8.00	92.57	8.6e+03	180725	AC104299	Homo sapiens chr
gb_hhg:AC026180	+	8.00	93.16	8.0e+03	166352	AC026180	Homo sapiens chrom	gb_hhg:AC103857	+	8.00	92.55	8.6e+03	181119	AC026180	Rattus norvegicus
gb_hhg:AC058792	+	8.00	93.16	8.0e+03	166772	AC058792	Homo sapiens chrom	gb_hhg:AC0092250	-	8.00	92.54	8.6e+03	181119	AC0092250	Homo sapiens chr
gb_hhg:AC022954	-	8.00	93.14	8.0e+03	167026	AP002783	Homo sapiens genom	gb_pr:AF411058	-	8.00	92.53	8.6e+03	181525	AF411058	Homo sapiens clo
gb_pr:AP002783	+	8.00	93.13	8.0e+03	167026	AP002783	Homo sapiens genom	gb_hhg:AC097538	+	8.00	92.53	8.6e+03	181525	AC097538	Rattus norvegicus
gb_hhg:AC000750	+	8.00	93.11	8.0e+03	167498	AC000750	Homo sapiens chrom	gb_pr:AC097714	+	8.00	92.53	8.6e+03	181756	AC097714	Homo sapiens chr
gb_hhg:AC103484	-	8.00	93.10	8.0e+03	167600	AC095510	Rattus norvegicus c	gb_hhg:AC027623	+	8.00	92.52	8.6e+03	181904	AC027623	Homo sapiens chr
gb_hhg:AC025094	+	8.00	93.07	8.0e+03	166481	AC025094	Homo sapiens chrom	gb_hhg:AC103954	+	8.00	92.52	8.6e+03	182022	AC103954	Homo sapiens chr
gb_hhg:AC058792	-	8.00	93.06	8.0e+03	168542	AC058792	Homo sapiens chrom	gb_hhg:AL535872	+	8.00	92.52	8.6e+03	182042	AL535872	Homo sapiens chr
gb_hhg:AC022954	-	8.00	93.06	8.0e+03	169370	AC022954	Homo sapiens clone	gb_hhg:AC021128	+	8.00	92.51	8.6e+03	182206	AC021128	Homo sapiens chr
gb_pr:AC0210854	+	8.00	93.02	8.1e+03	169682	AC010854	Homo sapiens chrom	gb_hhg:AC021128	+	8.00	92.51	8.6e+03	182206	AC021128	Homo sapiens chr
gb_hhg:AC0091147	+	8.00	93.01	8.1e+03	169836	AC091147	Homo sapiens clone	gb_pr:AL162578	+	8.00	92.48	8.7e+03	182823	AL162578	Homo sapiens chr
gb_hhg:AP0002435	+	8.00	93.00	8.1e+03	170028	AP002435	Homo sapiens chrom	gb_hhg:AC106545	+	8.00	92.46	8.7e+03	183553	AC106545	Rattus norvegicus
gb_pr:AC091970	+	8.00	93.00	8.1e+03	170113	AC091970	Homo sapiens chrom	gb_hhg:AC106545	+	8.00	92.45	8.7e+03	183646	AC106545	Rattus norvegicus
gb_hhg:AC018371	+	8.00	93.00	8.1e+03	170133	AC018371	Homo sapiens chrom	gb_pr:AL161646	+	8.00	92.45	8.7e+03	183833	AL161646	Human DNA sequen
gb_hhg:AC001753	+	8.00	92.99	8.1e+03	170204	AC018539	Homo sapiens chrom	gb_hhg:AC018708	+	8.00	92.44	8.7e+03	184516	AC018708	Homo sapiens chr
gb_pr:AP000773	+	8.00	92.99	8.1e+03	170434	AP000773	Homo sapiens genom	gb_pr:AC069242	+	8.00	92.41	8.7e+03	184664	AC069242	Homo sapiens 3 B
gb_hhg:AC0094970	+	8.00	92.98	8.1e+03	170682	AC094970	Rattus norvegicus c	gb_hhg:AC103508	+	8.00	92.41	8.7e+03	184794	AC103508	Homo sapiens chr
gb_hhg:AC023429	+	8.00	92.97	8.1e+03	170682	AC023429	Homo sapiens chrom	gb_hhg:AP000849	+	8.00	92.41	8.7e+03	184794	AP000849	Homo sapiens chr
gb_pr:AC087456	+	8.00	92.96	8.2e+03	170945	AC087456	Homo sapiens chrom	gb_pr:AC099314	+	8.00	92.40	8.8e+03	184945	AC099314	Homo sapiens chr
gb_hhg:AC025382	+	8.00	92.95	8.2e+03	171321	AC048382	Homo sapiens chrom	gb_pr:AC096641	-	8.00	92.39	8.8e+03	185148	AC096641	Homo sapiens chr
gb_hhg:AC003458	+	8.00	92.95	8.2e+03	171394	AC023287	Homo sapiens chrom	gb_hhg:AC006508	-	8.00	92.39	8.8e+03	185286	AC006508	Mus musculus Yp
gb_hhg:AC0091541	+	8.00	92.93	8.2e+03	171681	AC091541	Canis familiaris cl	gb_hhg:AL931723	+	8.00	92.37	8.8e+03	185644	AL931723	Homo sapiens chr
gb_pr:AC093278	+	8.00	92.93	8.2e+03	171858	AC093278	Homo sapiens chrom	gb_hhg:AC087163	+	8.00	92.37	8.8e+03	185680	AC087163	Homo sapiens chr
gb_hhg:AC090768	-	8.00	92.93	8.2e+03	171871	AC090768	Homo sapiens chrom	gb_hhg:AL645647	+	8.00	92.35	8.8e+03	186278	AL645647	Mus musculus chr
gb_hhg:AC0097419	-	8.00	92.93	8.2e+03	171890	AC097419	Rattus norvegicus c	gb_hhg:AL671978	-	8.00	92.34	8.8e+03	186497	AL671978	Homo sapiens chr
gb_hhg:AC001176	-	8.00	92.92	8.2e+03	172181	AC079816	Mus musculus clone	gb_hhg:AC011476	+	8.00	92.31	8.9e+03	187274	AC011476	Homo sapiens chr
gb_hhg:AC0010575	+	8.00	92.91	8.2e+03	172401	AC010575	Drosophila melanog	gb_in:AC007083	-	8.00	92.28	8.9e+03	188067	AC007083	Drosophila melan
gb_in:AC007467	+	8.00	92.90	8.2e+03	172401	AC010575	Drosophila melanog	gb_hhg:AL671978	-	8.00	92.26	8.9e+03	188067	AL671978	Mus musculus chr
gb_hhg:AC0099231	-	8.00	92.90	8.2e+03	172592	AC099231	Rattus norvegicus c	gb_pr:AL590410	+	8.00	92.26	8.9e+03	188615	AL590410	Human DNA sequen
gb_hhg:AC0098117	-	8.00	92.89	8.2e+03	172937	AC098117	Rattus norvegicus c	gb_hhg:AC011207	+	8.00	92.26	8.9e+03	188635	AC011207	Homo sapiens chr
gb_hhg:AC0025130	+	8.00	92.88	8.2e+03	172905	AC025130	Homo sapiens chrom	gb_hhg:AL670169	+	8.00	92.24	8.9e+03	189117	AL670169	Mus musculus chr
gb_hhg:AC003458	-	8.00	92.88	8.2e+03	173052	AP003458	Homo sapiens chrom	gb_pr:AC018992	+	8.00	92.17	9.0e+03	191055	AC018992	Homo sapiens BAC
gb_hhg:AC0034272	+	8.00	92.86	8.3e+03	173083	AC034272	Homo sapiens chrom	gb_hhg:AC069374	-	8.00	92.17	9.0e+03	191141	AC069374	Homo sapiens chr
gb_in:AC007467	+	8.00	92.86	8.3e+03	173422	AC007467	Drosophila melanog	gb_hhg:AC074358	-	8.00	92.14	9.1e+03	191775	AC074358	Mus musculus clo
gb_hhg:AC009520	+	8.00	92.86	8.3e+03	173479	AC097520	Homo sapiens chrom	gb_hhg:AC024731	-	8.00	92.12	9.1e+03	192493	AC024731	Homo sapiens chr
gb_hhg:AC021145	+	8.00	92.86	8.3e+03	173484	AC021145	Homo sapiens chrom	gb_pr:AC017075	+	8.00	92.09	9.1e+03	193277	AC017075	Homo sapiens BAC
gb_p1:AC093568	+	8.00	92.85	8.3e+03	173644	AC093568	Oryza sativa chrom	gb_hhg:AL671671	+	8.00	92.04	9.2e+03	194486	AL671671	Homo sapiens chr
gb_p1:AC093568	-	8.00	92.85	8.3e+03	173644	AC093568	Oryza sativa chrom	gb_hhg:AP001275	+	8.00	92.04	9.2e+03	194604	AP001275	Homo sapiens chr
gb_hhg:AC0216785	-	8.00	92.82	8.3e+03	174365	AL626785	Mus musculus chrom	gb_hhg:AC073389	-	8.00	92.03	9.2e+03	194735	AC073389	Homo sapiens chr
gb_hhg:AC005737	-	8.00	92.82	8.3e+03	174365	AL626785	Mus musculus chrom	gb_hhg:AC032024	-	8.00	92.01	9.2e+03	195466	AC032024	Homo sapiens chr
gb_hhg:AC102971	-	8.00	92.82	8.3e+03	174365	AC102971	Rattus norvegicus c	gb_hhg:AC087732	-	8.00	92.00	9.2e+03	195599	AC087732	Homo sapiens chr
gb_pr:AP002342	+	8.00	92.82	8.3e+03	174548	AP002342	Homo sapiens genom	gb_hhg:AL535651	+	8.00	91.97	9.3e+03	196414	AL535651	Homo sapiens chr
gb_hhg:AC0023869	+	8.00	92.81	8.3e+03	174612	AC023869	Homo sapiens chrom	gb_pr:AL535825	+	8.00	91.89	9.4e+03	196794	AL535825	Human DNA sequen
gb_hhg:AC0216785	-	8.00	92.81	8.3e+03	174637	AC0216785	Homo sapiens chrom	gb_hhg:AL644147	+	8.00	91.88	9.4e+03	199008	AL644147	Mus musculus chr
gb_hhg:AC0016896	-	8.00	92.81	8.3e+03	174742	AC016896	Homo sapiens chrom	gb_hhg:AL644147	+	8.00	91.88	9.4e+03	199008	AL644147	Yersinia pestis
gb_hhg:AC0091824	-	8.00	92.79	8.3e+03	175071	AC010274	Homo sapiens chrom	gb_hhg:AC003072	+	8.00	91.87	9.4e+03	199261	AC003072	Homo sapiens chr
gb_hhg:AC0097821	-	8.00	92.77	8.3e+03	175218	AC097821	Rattus norvegicus c	gb_p1:ATC8R145	+	8.00	91.86	9.4e+03	199548	ATC8R145	Arabidopsis thal
gb_hhg:AC0106567	+	8.00	92.77	8.4e+03	175660	AC106567	Rattus norvegicus c	gb_hhg:AC095457	+	8.00	91.85	9.4e+03	199708	AC095457	Rattus norvegicus
gb_hhg:AC0095485	+	8.00	92.76	8.4e+03	175837	AC095485	Rattus norvegicus c	gb_p1:ATF9A7	+	8.00	91.77	9.5e+03	201471	ATF9A7	Arabidopsis thal
gb_hhg:AC0091100	-	8.00	92.76	8.4e+03	175944	AC091									

gb.com:AF162514	+	7.00	115.65	444.57	462	AF162514 Bos grunniens satellite	gb.pr:HSN325321	-	7.00	111.91	718.00	780	AJ325321 Homo sapiens genomic
gb.pl:TL047411	-	7.00	115.61	446.33	464	U47411 Testamun labialis histon	em.hito.inh:AC056122	-	7.00	111.91	718.00	780	AC056122 Giardia intestinalis
gb.pr:HSR9GV9A	-	7.00	115.48	454.25	473	X05504 Human rearranged gene TR	em.hito.inh:AC074653	-	7.00	111.88	719.69	782	AC074653 Giardia intestinalis
gb.pr:MUSGBPA	-	7.00	115.46	455.13	474	X08086 Human rearranged gene TR	gb.pr:HSN335561	-	7.00	111.88	720.53	783	AJ335561 Homo sapiens genomic
gb.sts:HS006611	+	7.00	115.44	458.64	478	M57470 Murine beta-galactoside	em.hito.inh:AC030174	-	7.00	111.73	734.83	800	AC030174 Giardia intestinalis
gb.pl:IDV047383	-	7.00	115.15	472.68	489	U47383 Dumasia villosa histone	em.hito.inh:AC030179	-	7.00	111.63	744.08	811	AC081799 Giardia intestinalis
gb.ro:MM44KDEE	-	7.00	115.15	473.55	495	X53067 Mouse mRNA for 14kDa lec	gb.pat:AJ122305	+	7.00	111.59	748.27	816	AJ122305 Sequence 2221 from
gb.ro:MM6GLEEC	-	7.00	115.15	473.55	495	X53067 Mouse mRNA for 14kDa lec	gb.pl:YSPDC23A	+	7.00	111.53	753.31	822	L07304 Schistosoma genome
gb.pat:AJ27894	-	7.00	115.12	475.30	497	X27894 Coding sequence for GGP	em.hito.inh:AC071072	+	7.00	111.51	755.82	825	AC071072 Giardia intestinalis
gb.pat:AN111662	-	7.00	115.12	475.30	497	AR111662 Sequence 17 from paten	em.hito.inh:AC050107	+	7.00	111.45	761.69	832	AC050107 Giardia intestinalis
gb.sts:AF046519	-	7.00	115.11	476.18	498	AF046519 Rattus norvegicus, OTS	em.hito.inh:AC089928	+	7.00	111.35	771.74	844	AC089928 Giardia intestinalis
gb.com:AF045515	+	7.00	115.10	477.05	499	AF045515 Sus scrofa breed large	em.hito.inh:AC034448	+	7.00	111.33	773.41	846	AC034448 Giardia intestinalis
gb.ba:AF114969	-	7.00	114.86	481.43	504	AB039617 Chlorella sp. HS-7	em.hito.inh:AC071100	-	7.00	111.29	776.76	850	AC071100 Giardia intestinalis
gb.pl:AF114969	-	7.00	114.86	481.43	504	AB039617 Chlorella sp. HS-7	gb.pr:HSR9GV9A	-	7.00	111.28	778.43	852	X27894 Rattus norvegicus
gb.ro:AF095934	-	7.00	114.67	504.11	530	AF095934 Dictyostellum discoide	gb.pl:JUN1278454	-	7.00	111.24	781.78	856	AJ278454 Juglans nigra mRNA
gb.pr:HSR9GV9A	-	7.00	114.67	504.11	530	AF095934 Dictyostellum discoide	gb.pat:AB5851	+	7.00	111.12	794.31	871	AB5851 Sequence 510 from Pat
gb.pr:HSR9GV9A	-	7.00	114.65	504.11	530	AF095934 Dictyostellum discoide	gb.pat:AR155344	+	7.00	111.12	794.31	871	AR155344 Sequence 510 from Pat
gb.pr:HSR9GV9A	-	7.00	114.61	507.59	534	AR070588 Sequence 19 from paten	em.hito.inh:AC031263	+	7.00	111.12	794.31	871	AC031263 Giardia intestinalis
gb.pr:HSR9GV9A	-	7.00	114.59	509.33	536	277942 H. sapiens flow-sorted ch	em.hito.inh:AC044285	+	7.00	111.09	797.65	875	AC044285 Giardia intestinalis
gb.ro:BC002063	-	7.00	114.47	511.94	539	BC002063 Mus musculus, lectin	em.hito.inh:AC055802	+	7.00	111.09	808.45	888	AC055802 Sequence 1 from Pat
gb.ro:MM141L	-	7.00	114.47	511.94	539	BC002063 Mus musculus, lectin	gb.pr:BC007282	+	7.00	110.97	809.32	889	BC007282 Homo sapiens, clone
gb.pl:SOBETUA	-	7.00	114.43	519.76	548	L47262 Sportidololus paratensis	em.hito.inh:AC028084	+	7.00	110.87	820.14	902	AF222424 Plagiotecton suaveo
gb.ov:AF297036	+	7.00	114.36	524.10	553	AF297036 Ophiophagus hannah ubi	em.hito.inh:AC058409	+	7.00	110.85	822.64	911	AC058409 Giardia intestinalis
gb.pr:AB069580	-	7.00	114.20	535.36	566	M63999 Trilicium aestivum DNA b1	em.hito.inh:AC058409	+	7.00	110.85	822.64	911	AC058409 Giardia intestinalis
gb.pr:WHTEG619	-	7.00	114.18	537.96	569	M63999 Trilicium aestivum DNA b1	gb.sts:AB030249	+	7.00	110.80	827.63	912	AB030249 Synthetic Mus muscu
gb.pr:HSN33182	+	7.00	114.16	537.96	569	M63999 Trilicium aestivum DNA b1	em.hito.inh:AC053782	+	7.00	110.78	830.12	914	AC053782 Giardia intestinalis
gb.sts:G56256	+	7.00	114.15	538.82	570	G56256 SHGC-101770 Human Homo	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.in:AF165428	-	7.00	114.03	546.60	579	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.pat:AF330046	-	7.00	114.03	546.60	579	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.pat:AF330046	-	7.00	114.03	546.60	579	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo	em.hito.inh:AC058409	+	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	+	7.00	114.01	548.33	581	G58481 SHGC-104443 Human Homo							

gb_vt:EHU43560	-	7.00	109.08	1.0e+03	1159	U43560 Epizootic hemorrhagic d	gb_pat:AX061643	-	7.00	106.68	1.4e+03	1622	AX061643 Sequence 48 from P
gb_vt:AB002869	+	7.00	109.06	1.0e+03	1162	AB002869 Subacute sclerosing m	gb_vt:OYE22223	-	7.00	106.67	1.4e+03	1625	AJ292223 Onion yellow dwarf
gb_vt:AB016239	-	7.00	109.06	1.0e+03	1162	AB016239 Measles virus gene fo	gb_ov:AF124511	-	7.00	106.65	1.4e+03	1629	AF124511 Gallus gallus BVES
gb_pr:HSRKG9F	-	7.00	108.95	1.0e+03	1180	X15274 H.sapiens TRG9 gene, a	gb_sts:G73185	-	7.00	106.65	1.4e+03	1629	G73185 csprad1-pcr2-4 Huma
gb_pr:HSRKG9F	-	7.00	108.95	1.0e+03	1183	AJ27914 H.sapiens partial	gb_pr:AB070034	-	7.00	106.60	1.4e+03	1640	AB070034 Macaca fasciculari
gb_pr:BC008032	+	7.00	108.92	1.1e+03	1185	BC008032 Homo sapiens, clone I	gb_pr:BC011575	-	7.00	106.60	1.4e+03	1641	BC011575 Homo sapiens, Simi
gb_in:HSU29441	+	7.00	108.92	1.1e+03	1189	U29441 Diplomonad ATCC50330 be	gb_pl:PSALCBEH1	+	7.00	106.60	1.4e+03	1642	Y13338 Pichia stipitis ADH1
gb_pl:AY051038	-	7.00	108.85	1.1e+03	1198	AY051038 Arabidopsis thaliana	gb_pr:BC0282874	+	7.00	106.55	1.4e+03	1653	BC0282874 Homo sapiens, seri
gb_ba:PAR27709	-	7.00	108.85	1.1e+03	1200	U27709 Pseudomonas aeruginosa	gb_pl:AF370135	-	7.00	106.54	1.4e+03	1655	AF370135 Arabidopsis thalia
gb_in:EPHEPAD	-	7.00	108.82	1.1e+03	1203	D49925 Ephydratia fluviatilis m	gb_pl:AY042804	-	7.00	106.52	1.4e+03	1659	AY042804 Arabidopsis thalia
gb_pat:AX0083327	-	7.00	108.73	1.1e+03	1218	AX0083327 Sequence 19 from Pat	gb_pat:AX06562	+	7.00	106.51	1.4e+03	1662	AX06562 Sequence 595 from Pa
gb_pat:AX0029386	+	7.00	108.57	1.1e+03	1245	AX0029386 Sequence 5 from Paten	gb_pat:AX144962	-	7.00	106.48	1.4e+03	1669	AX144962 Sequence 8 from Pa
gb_in:AY032515	-	7.00	108.55	1.1e+03	1249	AF135155 Lutzomyia longipalpis	gb_pat:AX050448	-	7.00	106.48	1.4e+03	1669	AX050448 Sequence 8 from Pa
gb_in:AY069764	-	7.00	108.51	1.1e+03	1256	AY069764 Drosophila melanogast	gb_pat:AX336106	-	7.00	106.48	1.4e+03	1677	AX336106 Sequence 6615 from
gb_pl:LE5297916	+	7.00	108.48	1.1e+03	1261	AJ297916 Lycopersicon esculent	gb_pat:AX336460	-	7.00	106.48	1.4e+03	1677	AX336460 Sequence 6969 from
gb_vt:HX118854	+	7.00	108.43	1.1e+03	1270	Y18854 Infectious hematopoiet	gb_pr:HMUSHTA	-	7.00	106.48	1.4e+03	1679	HMUSHTA
gb_pl:AF139500	+	7.00	108.37	1.1e+03	1280	AF139500 Prunus americana p9f	gb_pat:AX030175	-	7.00	106.46	1.4e+03	1679	AX030175
gb_pr:AK05285	-	7.00	108.37	1.1e+03	1281	AK05285 Homo sapiens, cDNA FLJ	gb_pat:AX06584	+	7.00	106.45	1.4e+03	1677	AX06584 Sequence 597 from Pa
gb_pr:BC013062	-	7.00	108.36	1.1e+03	1282	BC013062 Homo sapiens, Similia	gb_pat:AX06584	+	7.00	106.45	1.4e+03	1677	AX06584 Sequence 601 from Pa
gb_ro:MUSELPL2	+	7.00	108.33	1.1e+03	1288	D49682 Mouse FT2-F1 gene, 2/19	gb_ro:MUSADPP	-	7.00	106.43	1.4e+03	1680	M93375 Mouse adipose differ
gb_ro:AF000204	+	7.00	108.29	1.1e+03	1295	AF000204 Rattus musculus ADP-rib	gb_pat:AX149469	-	7.00	106.42	1.5e+03	1683	AX149469 Sequence 25 from P
gb_ro:RN1LMY11	-	7.00	108.28	1.1e+03	1297	AF000204 Rattus sp. T-612 ret	gb_pat:AX123946	-	7.00	106.41	1.5e+03	1685	AX123946 Sequence 1 from pat
gb_pat:153503	-	7.00	108.28	1.1e+03	1297	U87600 Rattus norvegicus L1 re	gb_pat:AX051970	-	7.00	106.39	1.5e+03	1690	AX051970 Sequence 24 from p
gb_ro:AF000199	-	7.00	108.27	1.1e+03	1299	163503 Sequence 3 from patent	gb_pat:152144	-	7.00	106.39	1.5e+03	1690	152144 Sequence 24 from pat
gb_pr:HS11P0PBL	+	7.00	108.23	1.2e+03	1306	AF000199 Rattus sp. T-612 ret	gb_pr:AK000500	+	7.00	106.39	1.5e+03	1690	AK000500 Homo sapiens cDNA
gb_pr:AF43203	+	7.00	108.19	1.2e+03	1314	X97324 H.sapiens mRNA for adif	gb_pr:AF42529	-	7.00	106.37	1.5e+03	1695	AF42529 Homo sapiens prote
gb_pl:AY038325	-	7.00	108.13	1.2e+03	1335	AY038325 Homo sapiens adipose	gb_pl:ST79359	-	7.00	106.35	1.5e+03	1700	ST79359 arabidogalactan prot
gb_pl:ATM271731	-	7.00	108.09	1.2e+03	1335	AY038325 Inocbe sp. Triappe 25	gb_vt:AF226646	-	7.00	106.29	1.5e+03	1714	AF226646 Turnip mosaic viru
gb_pat:AX149467	+	7.00	108.01	1.2e+03	1347	AJ271731 Arabidopsis thaliana	gb_pl:DC047048	-	7.00	106.27	1.5e+03	1719	DC047048
gb_in:AY061465	+	7.00	107.98	1.2e+03	1352	184655 Sequence 10 from patent	gb_hlg:AC106945	-	7.00	106.25	1.5e+03	1723	AC106945 Rattus norvegicus
gb_in:AY061465	+	7.00	107.87	1.2e+03	1353	AY061465 Sequence 10 from patent	gb_ro:AB049055	-	7.00	106.23	1.5e+03	1736	AB049055 Mus musculus Brail
gb_pr:HSK32592	+	7.00	107.85	1.2e+03	1376	AY065455 Drosophila melanogast	gb_pl:AB049055	-	7.00	106.23	1.5e+03	1736	AB049055 Mus musculus Brail
gb_vt:AM80056	+	7.00	107.84	1.2e+03	1380	AJ325792 Homo sapiens genomic	gb_ov:AF197048	-	7.00	106.01	1.5e+03	1782	AF197048 Sperm aurata para
gb_pl:WHREMGFE13	+	7.00	107.73	1.2e+03	1406	M62893 Trifolium aestivum DNA H	gb_pl:HMV31084	-	7.00	105.99	1.5e+03	1787	HMV31084 Sperm aurata para
gb_ro:AB049056	+	7.00	107.70	1.2e+03	1406	AB049056 Rattus norvegicus bra	gb_pl:AB049056	-	7.00	105.99	1.5e+03	1787	AB049056 Rattus norvegicus
gb_ba:AB050224	+	7.00	107.70	1.2e+03	1407	AB050224 Uncultured archaean S	gb_pat:AX086525	-	7.00	105.75	1.6e+03	1848	AX086525 Sequence 475 from
gb_pat:119715	-	7.00	107.67	1.2e+03	1412	119715 Sequence 1 from patent	gb_pr:HSR801723	-	7.00	105.72	1.6e+03	1857	HSR801723
gb_ro:MMU65021	-	7.00	107.67	1.2e+03	1413	AF039461 Mus musculus Leptin r	gb_ro:MMU65021	+	7.00	105.71	1.6e+03	1860	MMU65021 Mus musculus amelobl
gb_da:TRHEMCA	-	7.00	107.65	1.2e+03	1416	L20680 Thermus aquaticus DNA r	gb_pr:HS11485	-	7.00	105.68	1.6e+03	1867	HS11485 Homo sapiens mRNA fo
gb_ro:RNM1LTV11	-	7.00	107.55	1.2e+03	1429	U87598 Rattus norvegicus L1 re	gb_ro:BC006906	-	7.00	105.64	1.6e+03	1877	BC006906 Mus musculus, Simi
gb_ro:AF047384	+	7.00	107.56	1.3e+03	1435	AF047384 Rattus norvegicus pos	gb_pat:AX068976	-	7.00	105.55	1.6e+03	1890	AX068976 Sequence 1 from Pa
gb_pr:U398980	-	7.00	107.51	1.3e+03	1445	U398980 Human PM2 related (PM	gb_in:AY035782	-	7.00	105.55	1.6e+03	1900	AY035782 Rhinophallus appe
gb_pl:TA0019733	-	7.00	107.25	1.3e+03	1499	U07933 Trifolium aestivum Chind	gb_pat:AX001278	+	7.00	105.55	1.6e+03	1900	AX001278 Sequence 22 from p
gb_ro:MMU61282	-	7.00	107.23	1.3e+03	1502	U49107 Mus musculus Leptin rec	gb_pat:AX025098	-	7.00	105.55	1.6e+03	1900	AX025098 Sequence 22 from p
gb_in:AY061282	-	7.00	107.17	1.3e+03	1515	AY061282 Drosophila melanogast	gb_ro:RNM21070	-	7.00	105.55	1.6e+03	1900	RNM21070
gb_pat:AX353444	+	7.00	107.16	1.3e+03	1518	AX353444 Sequence 10 from Pat	gb_ba:AF056335	-	7.00	105.54	1.6e+03	1905	AF056335 Bacillus lichenifo
gb_pat:144674	-	7.00	107.15	1.3e+03	1520	144674 Sequence 1 from patent	gb_in:AK026783	+	7.00	105.52	1.6e+03	1914	AK026783 Homo sapiens cDNA:
gb_ro:RNEHBAB	-	7.00	107.13	1.3e+03	1524	Z28365 R.norvegicus BEHA mRNA	gb_in:AX058081	-	7.00	105.50	1.6e+03	1914	AX058081 Bombyx mori putati
gb_pl:AY054265	-	7.00	107.13	1.3e+03	1524	AY054265 Arabidopsis thaliana	gb_pr:BC005127	+	7.00	105.49	1.6e+03	1916	BC005127 Homo sapiens, adip
gb_pr:AB00161S02	-	7.00	107.13	1.3e+03	1524	AB001616 Homo sapiens DNA for	gb_da:TVCPPT	-	7.00	105.47	1.6e+03	1923	TVCPPT
gb_vt:AF448803	-	7.00	107.09	1.3e+03	1532	AF448803 Human herpesvirus 4 E	gb_vt:AF448803	-	7.00	105.47	1.6e+03	1923	AF448803 Human herpesvirus 4 E
gb_pat:AR023708	-	7.00	107.08	1.3e+03	1535	AR023708 Sequence 10 from pat	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pat:AR103342	-	7.00	107.01	1.3e+03	1550	AR103342 Sequence 3 from paten	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pat:BD000105	-	7.00	107.01	1.3e+03	1550	BD000105 Differentiation and t	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pr:HS058681	-	7.00	106.99	1.3e+03	1550	U58681 Homo sapiens neurogenic	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pl:AY065116	+	7.00	106.97	1.4e+03	1559	AY065116 Arabidopsis thaliana	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_in:AY060613	-	7.00	106.97	1.4e+03	1559	AY060613 Drosophila melanogast	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pat:AX136103	-	7.00	106.85	1.4e+03	1584	AX136103 Sequence 25 from Pat	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pat:AX349757	-	7.00	106.85	1.4e+03	1584	AX349757 Sequence 280 from Pat	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pr:AF038193	+	7.00	106.84	1.4e+03	1585	AF038193 Homo sapiens clone 23	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pr:HS14487	-	7.00	106.84	1.4e+03	1586	Y14487 Homo sapiens mRNA for c	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pr:HMUSHTA	-	7.00	106.79	1.4e+03	1586	Y14487 Homo sapiens mRNA for c	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_in:AA03190	+	7.00	106.77	1.4e+03	1602	BC018757 Homo sapiens, clone M	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pr:BC018757	-	7.00	106.77	1.4e+03	1603	BC018757 Homo sapiens, clone M	gb_vt:IHNGNVB	+	7.00	105.43	1.6e+03	1933	IHNGNVB
gb_pat:AX083337	-	7.00	106.75	1.4e+03	1606	AX083337 Sequence 29 from Pat	gb_pr:AF076619	-	7.00	105.37	1.7e+03	1930	AF076619 Rattus norvegicus
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:AX036113	-	7.00	105.25	1.7e+03	1982	AX036113 Sequence 12 from P
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_pat:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:IHNS0401	+	7											

gb_da:SCU60175	-	7.00	105.25	1.7e+03	1984	U60175 Sphingomonas chlorophen	gb_vi:SQLCVE	-	7.00	103.30	2.2e+03	2607	M38182 Squash leaf curl v1r
gb_ro:RNLMLV1	-	7.00	105.24	1.7e+03	1986	U87602 Rattus norvegicus L1 re	gb_pat:AX36123	-	7.00	103.19	2.2e+03	2645	AX036123 Sequence 22 from p
gb_pat:RNLMLV1	+	7.00	105.23	1.7e+03	1987	AX281642 Sequence 51 from Pat	gb_pat:158538	-	7.00	103.19	2.2e+03	2645	158538 Sequence 22 from pat
gb_da:PEP0133737	+	7.00	105.19	1.7e+03	2000	AI133737 Paenibacillus polymy	gb_pat:187013	-	7.00	103.19	2.2e+03	2645	187013 Sequence 22 from pat
gb_pl:AY0406011	-	7.00	105.16	1.7e+03	2009	AY046011 Arabidopsis thaliana	gb_ro:BC003668	-	7.00	103.19	2.2e+03	2645	BC003668 Homo sapiens, clon
gb_pr:BC022461	-	7.00	105.12	1.7e+03	2019	BC022461 Homo sapiens, neurog	gb_ro:BC021440	-	7.00	103.16	2.2e+03	2658	BC021440 Homo musculus, clon
gb_da:EA34346	-	7.00	105.10	1.7e+03	2026	E34346 DNA and process for pcc	gb_pr:AF276889	-	7.00	103.13	2.2e+03	2671	AF276889 Homo sapiens amaro
gb_in:AF216210	+	7.00	105.01	1.7e+03	2050	AF216210 Drosophila buzzatii	gb_pl:AB037203	-	7.00	103.12	2.2e+03	2671	AB037203 Glycyrhiza glabra
gb_pl:AY050818	+	7.00	105.00	1.7e+03	2054	AY050818 Arabidopsis thaliana	gb_da:BC523232	-	7.00	103.11	2.2e+03	2676	AF352392 Candidatus Procoba
gb_ro:MUSC0142	+	7.00	105.00	1.7e+03	2054	D38412 Mouse mRNA for pro-aldh	gb_ro:BC005672	-	7.00	103.11	2.2e+03	2677	BC005672 Mus musculus, clon
gb_pat:AX036112	-	7.00	104.98	1.7e+03	2060	AX036112 Sequence 11 from Pat	gb_da:AF062363	-	7.00	103.10	2.2e+03	2678	AF352393 Candidatus Procoba
gb_pat:1585027	-	7.00	104.98	1.7e+03	2060	158527 Sequence 11 from patent	gb_ov:IRU068757	+	7.00	103.10	2.2e+03	2680	U87857 Rana ridibunda secre
gb_pat:1585027	-	7.00	104.98	1.7e+03	2060	158527 Sequence 11 from patent	gb_ov:XLSG1	+	7.00	103.06	2.2e+03	2696	X32873 x laevis mRNA for se
gb_pl:AF418308	+	7.00	104.91	1.7e+03	2064	AF418308 Arabidopsis thaliana	gb_pr:AF038192	-	7.00	103.05	2.2e+03	2700	AF008192 Homo sapiens putat
gb_ro:BC013564	+	7.00	104.91	1.7e+03	2080	BC013564 Mus musculus, RIKEN	gb_in:DMNRS	+	7.00	103.04	2.2e+03	2708	U94513 E. melanoagaster mRNA
gb_pl:SCYU01033C	+	7.00	104.88	1.8e+03	2089	Z74081 S. cerevisiae chromosome	gb_da:094413	-	7.00	103.03	2.2e+03	2708	U94513 E. melanoagaster mRNA
gb_pl:LES193372	+	7.00	104.79	1.8e+03	2115	AI133732 Arabidopsis thaliana	gb_in:AY0404294	-	7.00	102.98	2.3e+03	2726	AY0404294 Alveolinella quol
gb_pl:LES193372	+	7.00	104.77	1.8e+03	2122	AI306630 Lycopersicon esculent	gb_da:SPU6330	-	7.00	102.98	2.3e+03	2784	U36390 Synechococcus PCC700
gb_pr:HSAPNPK1	-	7.00	104.75	1.8e+03	2126	AI306630 Human Na ⁺ , K ⁺ ATPase gen	gb_da:RSEKXOYG	-	7.00	102.79	2.3e+03	2800	X16704 Rhizobium sp. exoA
gb_pl:AF306764	-	7.00	104.70	1.8e+03	2143	AF306764 Cochliobolus carbonum	gb_da:BFU05415	-	7.00	102.74	2.3e+03	2819	U64515 Bacillus firmus puta
gb_pr:AF258591	-	7.00	104.69	1.8e+03	2145	AF258591 Homo sapiens p784 m	gb_pr:AF355121	-	7.00	102.72	2.3e+03	2826	AF355121 Pan troglodytes cl
gb_pat:ES2146	-	7.00	104.60	1.8e+03	2172	ES2146 Novel protein and DNA	gb_pl:AF467983	-	7.00	102.71	2.3e+03	2830	AF467983 Pan troglodytes cl
gb_da:ES2150	-	7.00	104.60	1.8e+03	2172	ES2150 Novel protein and DNA	gb_pr:AF355118	-	7.00	102.70	2.3e+03	2835	AF355118 Pan troglodytes cl
gb_pr:AF270549	+	7.00	104.53	1.8e+03	2193	AE270549 Homo sapiens clone 18	gb_pr:AF355126	-	7.00	102.69	2.3e+03	2838	AF355126 Homo sapiens clone
gb_pr:HSAC0145	+	7.00	104.51	1.9e+03	2199	U41068 Human collagen alpha2(X	gb_pat:AB3691	-	7.00	102.69	2.3e+03	2838	AF3691 Sequence 18 from Pat
gb_pat:AX353714	+	7.00	104.51	1.9e+03	2200	AX353714 Sequence 25 from Pat	gb_vi:AB022887	-	7.00	102.67	2.3e+03	2845	AF355117 Pan troglodytes cl
gb_pr:AB047835	-	7.00	104.46	1.9e+03	2215	AB047835 Macaca fascicularis	gb_pr:AF355120	-	7.00	102.67	2.3e+03	2847	AF355120 Pan troglodytes cl
gb_da:SYO518	-	7.00	104.44	1.9e+03	2220	M8825 Synechococcus sp. Isla	gb_pr:AB072792	-	7.00	102.65	2.4e+03	2853	AB072792 Macaca fasciculari
gb_pr:HUMMEL18	+	7.00	104.42	1.9e+03	2227	AI3969 Human mRNA for Mel-18	gb_pr:AB022202	-	7.00	102.65	2.4e+03	2863	AB022202 Macaca fasciculari
gb_da:VIBMETRA	-	7.00	104.41	1.9e+03	2232	X68658 Vibrio sp. gene for V	gb_da:BP335	-	7.00	102.63	2.4e+03	2864	X68321 B. pertussis gene for
gb_pl:VEVCINMTR	-	7.00	104.37	1.9e+03	2240	X49831 V. faba VEVCINV mRNA for	gb_da:BP335	-	7.00	102.63	2.4e+03	2865	X68321 B. pertussis gene for
gb_com:OCU8153	+	7.00	104.29	1.9e+03	2267	U38193 Oryctolagus cuniculus	gb_da:BP335	-	7.00	102.62	2.4e+03	2865	X68321 B. pertussis gene for
gb_ro:AF024620	-	7.00	104.23	1.9e+03	2286	AF024620 Mus musculus gamma-am	gb_da:BP335	-	7.00	102.57	2.4e+03	2885	X68321 B. pertussis gene for
gb_ro:RNLMAF	-	7.00	104.23	1.9e+03	2288	X06942 Rat A-raf mRNA, 9/1/93	gb_da:BP335	-	7.00	102.56	2.4e+03	2888	X68321 B. pertussis gene for
gb_da:BACOR160	-	7.00	104.23	1.9e+03	2288	X07687 Rat L1Rn B7 repetitive	gb_ro:MUSC033003	-	7.00	102.56	2.4e+03	2892	X22574 M. musculus gamma-C-C
gb_pat:AF015229	-	7.00	104.22	1.9e+03	2290	M60475 B. thuringiensis replicat	gb_da:BP335	-	7.00	102.55	2.4e+03	2896	X68321 B. pertussis gene for
gb_pl:AF115272	-	7.00	104.21	1.9e+03	2290	AF115272 Sequence 1 from patent	gb_pr:AF355126	-	7.00	102.52	2.4e+03	2899	AF355126 Homo sapiens p47-pho
gb_pr:AF448510	-	7.00	104.06	2.0e+03	2344	AF448510 Homo sapiens histone	gb_com:AB008467	-	7.00	102.52	2.4e+03	2908	AF355117 Pan troglodytes cl
gb_in:ABR239060	+	7.00	104.05	2.0e+03	2347	AF332141 Mus musculus receptor	gb_in:AY070992	-	7.00	102.49	2.4e+03	2917	AY070992 Drosophila melanog
gb_pl:AB026830	+	7.00	103.97	2.0e+03	2373	AI239060 Trypanosoma brucei B5	gb_pl:AFH92782	-	7.00	102.42	2.4e+03	2919	AF297282 Arabidopsis thaliana
gb_pr:HUMORB14R	-	7.00	103.96	2.0e+03	2376	L0266890 Nicotiana tabacum mRN	gb_pr:U0MAPE500	-	7.00	102.40	2.4e+03	2956	M16446 Human adenine phosph
gb_pl:AB063055	-	7.00	103.96	2.0e+03	2376	L16687 Homo sapiens Gb14 mRN	gb_ro:BC016500	-	7.00	102.29	2.5e+03	3002	BC016500 Mus musculus, inte
gb_pl:AB063055	-	7.00	103.96	2.0e+03	2376	AF000515 Homo sapiens cDNA FR1	gb_in:AY071210	-	7.00	102.28	2.5e+03	3005	AY071210 Drosophila melanog
gb_da:AF092006	-	7.00	103.92	2.0e+03	2389	AF092006 Bacillus licheniform	gb_pr:HSAPRT	-	7.00	102.26	2.5e+03	3016	AF092006 Homo sapiens gene Ap
gb_in:AF092006	+	7.00	103.88	2.0e+03	2404	AF238309 Drosophila melanogast	gb_ro:AF033919	-	7.00	102.22	2.5e+03	3030	AF023919 Mus musculus PK-12
gb_ov:AF098788	+	7.00	103.88	2.0e+03	2404	AF098788 Gallus gallus nucleat	gb_pat:AX058011	-	7.00	102.21	2.5e+03	3034	AX058011 Homo sapiens cDNA
gb_da:AF079316	+	7.00	103.87	2.0e+03	2407	AF079316 Shigella flexneri str	gb_pat:AX058011	-	7.00	102.21	2.5e+03	3034	AX058011 Homo sapiens cDNA
gb_pr:BC007979	+	7.00	103.84	2.0e+03	2416	BC007979 Homo sapiens, serine	gb_vi:HSAP5E51	-	7.00	102.15	2.5e+03	3041	AF004865 Homo sapiens mRNA
gb_ro:BC021840	+	7.00	103.84	2.0e+03	2416	BC021840 Mus musculus, clone	gb_vi:HSAP5E51	-	7.00	102.15	2.5e+03	3060	M00868 Epstein-Barr virus 1
gb_ro:AF428262	-	7.00	103.81	2.0e+03	2427	AF428262 Mus musculus calcitri	gb_pr:AK001566	-	7.00	102.15	2.5e+03	3062	AK001566 Homo sapiens cDNA
gb_ro:AK023141	-	7.00	103.67	2.1e+03	2475	AK023141 Homo sapiens cDNA FLJ	gb_pat:AX029571	-	7.00	102.11	2.5e+03	3080	AX29571 Sequence 80 from p
gb_da:BMGLE	+	7.00	103.64	2.1e+03	2484	X37448 B. meliensis gale, mazi	gb_pr:AB004884	-	7.00	102.11	2.5e+03	3080	AB004884 Homo sapiens mRNA
gb_pat:ES2149	+	7.00	103.60	2.1e+03	2499	ES2149 Novel protein and DNA	gb_ro:AF033919	-	7.00	102.05	2.5e+03	3106	AF033919 Homo sapiens CGMP-
gb_in:AF149286	+	7.00	103.60	2.1e+03	2500	AF149286 Caenorhabditis elegans	gb_ro:AF033919	-	7.00	101.94	2.5e+03	3159	AF033919 Homo sapiens mRNA for
gb_pr:AK001615	+	7.00	103.59	2.1e+03	2503	AK001615 Homo sapiens cDNA FLJ	gb_pl:SF6927	-	7.00	101.93	2.5e+03	3159	AF033919 Homo sapiens mRNA for
gb_pl:PSA01312	+	7.00	103.58	2.1e+03	2506	AF223312 Pinus sylvestris mltc	gb_ro:AF033919	-	7.00	101.89	2.5e+03	3173	AF033919 Homo sapiens mRNA for
gb_pl:PSA01312	+	7.00	103.58	2.1e+03	2506	AF223312 Pinus sylvestris mltc	gb_ro:AF033919	-	7.00	101.89	2.5e+03	3173	AF033919 Homo sapiens mRNA for
gb_pat:AX149316	+	7.00	103.56	2.1e+03	2511	AX149316 Sequence 1 from Patent	gb_ro:AF033919	-	7.00	101.89	2.5e+03	3173	AF033919 Homo sapiens mRNA for
gb_hlg:AX098341	+	7.00	103.55	2.1e+03	2511	AX098341 Rattus norvegicus c1c	gb_ro:AF033919	-	7.00	101.89	2.5e+03	3173	AF033919 Homo sapiens mRNA for
gb_pl:LES03343	+	7.00	103.54	2.1e+03	2519	AF303343 Lycopersicon esculent	gb_in:DMO5BP	-	7.00	101.73	2.6e+03	3247	AF303343 Lycopersicon esculent
gb_pr:HSR000585	+	7.00	103.52	2.1e+03	2521	AF080084 Homo sapiens mRNA; cd	gb_pat:AX060472	-	7.00	101.68	2.7e+03	3267	AF060472 Sequence 12 from p
gb_da:STIMPABC	+	7.00	103.52	2.1e+03	2527	X33528 Salmonella typhimurium	gb_pr:AK023464	-	7.00	101.68	2.7e+03	3269	AK023464 Homo sapiens cDNA
gb_pr:AY0455919	+	7.00	103.48	2.1e+03	2527	AY0455919 Arabidopsis thaliana	gb_in:AF238308	-	7.00	101.66	2.7e+03	3327	AF238308 Drosophila melanog
gb_ov:AB019790	-	7.00	103.46	2.1e+03	2547	AB019790 Gallus gallus mRNA for	gb_pr:AF122667	-	7.00	101.56	2.7e+03	3327	AF122667 Homo sapiens tounsl
gb_da:AF227554	-	7.00	103.46	2.1e+03	2547	AF227554 Porphyromonas gingiv	gb_pr:HS47P4302	-	7.00	101.51	2.7e+03	3348	HS47P4302 Homo sapiens p47-pho
gb_pr:AK056588	-	7.00	103.44	2.1e+03	2556	AK056588 Homo sapiens cDNA FLJ	gb_pr:HS47P4302	-	7.00	101.51	2.7e+03	3350	HS47P4302 Homo sapiens p47-pho
gb_ov:ICSTSTSBP	-	7.00	103.37	2.1e+03	2580	X16078 T. californica mRNA for	gb_pat:AX024038	-	7.00	101.44	2.7e+03	3381	AX024038 Sequence 7 from Pa
gb_pr:BC018109	-	7.00	103.36	2.1e+03	2585	BC018109 Homo sapiens, Stimulat	gb_pl:SCYU142C	-	7.00	101.44	2.8e+03	3383	U49417 S. cerevisiae chromos

gb_pl:AF053127 - 7.00 101.39 2.8e+03 3405 ! AF053127 *Malus domestica* leuc
gb_to:AF334942 - 7.00 101.30 2.8e+03 3449 ! AF334942 *Mus musculus* toll-11k
gb_to:AF334943 - 7.00 101.27 2.8e+03 3463 ! AF334943 *Mus musculus* toll-11k
gb_ba:AF033314 - 7.00 101.23 2.8e+03 3483 ! AF033314 Uncultured gamma prot
gb_pat:AR129174 + 7.00 101.22 2.8e+03 3487 ! AR129174 Sequence 32 from pat
gb_to:MMKCOOBRB - 7.00 101.20 2.8e+03 3499 ! Y10296 *Mus musculus* mRNA for 1
gb_to:MMNZOOBRB - 7.00 101.20 2.8e+03 3499 ! Y10298 *Mus musculus* mRNA for 1
gb_in:AF245455 - 7.00 101.18 2.8e+03 3507 ! AF245455 *Drosophila melanogaster*
gb_ba:MG249993 - 7.00 101.13 2.9e+03 3530 ! AJ249993 *Mytilus galli* oprovinc
gb_ba:AY033326 - 7.00 101.10 2.9e+03 3546 ! AY033326 Uncultured gamma prot
gb_pl:AY033311 - 7.00 101.00 2.9e+03 3597 ! AY033311 Uncultured gamma prot
gb_pl:SCYAK1 + 7.00 101.00 2.9e+03 3599 ! X16056 *Yeast YAK1* gene for yak
gb_to:AF291054 + 7.00 100.96 2.9e+03 3617 ! AF291054 *Mus musculus* 8-oxodg

seq_name: gb_ba:AF242417

seq_documentation_block:
LOCUS AF242417 777 bp DNA linear BCT 01-JAN-2002
DEFINITION *Escherichia coli* heat-labile enterotoxin subunit A precursor, gene,
complete cds.
ACCESSION AF242417 GI:18026885
VERSION AF242417.1 GI:18026885
KEYWORDS
SOURCE *Escherichia coli*.
ORGANISM *Escherichia coli*.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 777)
AUTHORS Chang G.N. and Ho K.C.
TITLE Heat-labile enterotoxin subunit A gene of *Escherichia coli*
JOURNALS Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Chang G.N. and Ho K.C.
TITLE Direct Submission
JOURNALS Submitted (07-MAR-2000) Botany, National Taiwan University 1
Roosevelt Road, Sec. 4, Taipei, Taiwan 106, Republic of China
FEATURES
Source
1. 777
Location/Qualifiers
/organism="Escherichia coli"
/strain="CH-C2"
/serotype="O-136"
/specific_host="pig"
/db_xref="taxon:562"
/country="Taiwan"
/note="Isolated from hemolytic cells"
1. 777
/note="ltpa"
/codon_start=1
/transl_table=1
/product="heat-labile enterotoxin subunit A precursor"
/protein_id="AAU55671.1"
/db_xref="GI:18026886"
/translation="MKNTFFIFLLASPLYANGDRLYRADSRPDEIKRSGIMPRG
HNEYFDGTMQNTINLYATGAGTGTGPRYDGYVSTSLRSALHAGOSLSGYSITY
IYIATAPNENYNDVLCVSPHYEDVSALGIPYSQITGMTRVEGVIDELHKN
REYRDYRNLIAPADGRLAGPPDQAMKEEPIIHAPGCGSSRTIGDYCN
EETQNLSTIYLRVQSKVRQIFSDYQSEVDIYNRIDEL"
1. 54
55. 630
55. 630
/product="heat-labile enterotoxin subunit A1"
637. 774
/product="heat-labile enterotoxin subunit A2"
BASE COUNT 255 a 136 c 164 g 222 t
ORIGIN

alignment_scores:
Quality: 240.00 Length: 240
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AF242417 ..

Align seg 1/1 to: AF242417 from: 1 to: 777

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
55 AATGGCGACAGATTATACCGGTGCTGACTGTAGACCCCGATGAAATAA 104
17 sArSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
105 ACGTTCGCGAGGTGTTATAGCCAGAGGCGCATATAGACTTCGATGAG 154
34 lYThGlmMetAsnIleAsnLeuTyrAspHisAlaArgGlyThGlnThr 50
155 GAACCAATGAAATATATATCTTATGATCAGCGAGAGCAACCAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
205 GCGTTTGTACATATGATGATGATGATGATGATGATGATGATGATG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
255 AAGTCTCAGCTTAGCAGACAGCTCATATTTATCAGATATTCAGCTTACT 304
84 yrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
305 ATATATATGTTATATAGCAGACACCAATATGTTAATGTTAATGATG 354
101 leuGlyValIleTyrSerProHisProTyrGluGluGluValSerAlaLeuG 117
355 TTAGGCGTATACAGCCCTCCACCATATGACAGAGGTTCTTGCGTTAGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
405 TGGATACCATATATCTCAGATATATGATGATGATGATGATGATGATG 454
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
455 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
505 AGAAATCTGATATATGCTCCGCGAGAGATGTTACAGATTAAGACAGTTT 554
167 eProProAspHisGlnAlaTyrPArgGluGluProTyrIleHisAlaP 184
555 CCCACCGGATCACCAAGCTTGAGAGAAACCCGATTCATCATCGAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleHnGlyAspThrCysAsn 200
605 CACAAGGTTGTGGAATTCATCAAGAACCAATCACAGGTGATCTGTGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIle 217
655 GAGGAGACCCAGATCTAGACATATATCTCAGAGGATATCATCAATAA 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
705 AGTTAAGAGCGCAGATATTTTCACTATCAGAGTTCAGAGTTCACATATA 754
234 snArgIleArgAspGluLeu 240
755 ACAGAAATTCGGAGTGAATTA 774

seq_name: gb_ba:P307LTA

seq_documentation_block:
LOCUS P307LTA 777 bp DNA linear BCT 24-APR-1996
DEFINITION Plasmid P307 (from *E. coli*) heat-labile enterotoxin subunit A (lta)
gene, complete cds.
ACCESSION M35581
VERSION M35581.1 GI:150458
KEYWORDS enterotoxin; lta gene.
SOURCE Plasmid P307 (clone: PAT153.) DNA.

ORGANISM Plasmid P307
 plasmids.
 1 (bases 1 to 777)
 REFERENCE Dykes,C.W., Halliday,I.J., Hobden,A.N., Read,M.J. and Harford,S.
 AUTHORS A comparison of the nucleotide sequence of the A subunit of heat-
 TITLE labile enterotoxin and cholera toxin
 JOURNAL FEMS Microbiol. Lett. 26, 171-174 (1985)
 FEATURES
 source
 1..777
 /organism="Plasmid P307"
 /plasmid="Plasmid P307"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2472"
 /clone="PAT153."
 1..777
 /gene="lta"
 1..777
 /gene="lta"
 CDS
 1..777
 /gene="lta"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AA98202.1"
 /db_xref="GI:150459"
 /translation="MKNTFFIFILLASPLYANGDRLYRADSRPPEIKRSGGLMPRG
 HNEYDRGTOMINLYDHARGOTGFVARDGYSVTSLSRAHLAGOSILSGYSTY
 IYVATAPNMENVDLVGYSHPHYEOEVALGGIPLYQIIGWYRVNFGVIDERLHRN
 REYDRYRNLIAPAEDEGYRLAGPRPDQAREPPIHHAPOGCCNSSRTTGDTCN
 EETONLSTIYREYOSKVKROIYFSYOSEVDIYNIRDEL"
 BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN
 alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 us-09-528-682-1 x P307LTA ..
 Align seg 1/1 to: P307LTA from: 1 to: 777

134 ATTTAASPGIUAARGLEUHHISARGASNRGILUTYRARGASPARGTYYT 150
 455 TGAATGATGAACGATTACATGATACAGGAAATATAGACCGGATATAC 504
 151 ARGASNEUASNIIEALAPROALAGLUASPGIYTYRARGLEUALAGLYP 167
 505 AGAAATGTGAATATAGCTCCGGCAGAGAGATGGTTACGATTTACAGSTTT 554
 167 ePPROASPPIHISGLNALATPARGIUGIUPROTPTPIHHISALAP 184
 555 CCCACCCGATCACCACAGCTTGAGAGAAACCTTGATTCATCATCAGCAC 604
 184 roGINIGYSGIYASNSeSerSerArthrIleThrgIYAspPthrCYasn 200
 605 CACAGAGTTGTGAAATTCATCAGAACAAATCACAGGTGACTTGTAAAT 654
 201 GUGIUThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerly 217
 655 GAGGAGACCCAGAAATCTGACACAAATATATCTCAGGAAATATCATCAAA 704
 217 sVALYsARGIlnIlePheserAspTYrGlnSerGluValAspIleTYrA 234
 705 AGTTAAGAGCCAGATATTTTCAGACTATCATGACAGAGTTGACATATATA 754
 234 snArGIlleArgAspGluLeu 240
 755 ACAGAAATTCGGAGAAATTA 774
 seq_name: gb_pat:A04913
 seq_documentation_block:
 LOCUS A04913 777 bp DNA linear PAT 13-JUL-1993
 DEFINITION E.coli LTA gene for labile toxin A.
 ACCESSION A04913
 VERSION A04913.1 GI:412215
 KEYWORDS labile toxin A.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes,M.V., Harford,S. and Ross,G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 5 19-JUN-1985;
 GLAXO GROUP LIMITED
 FEATURES
 source
 1..777
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..777
 /codon_start=1
 /transl_table=11
 /product="labile toxin A"
 /protein_id="CAA00402.1"
 /db_xref="GI:412216"
 /db_xref="SWISS-PROT:P06717"
 /translation="MKNTFFIFILLASPLYANGDRLYRADSRPPEIKRSGGLMPRG
 HNEYDRGTOMINLYDHARGOTGFVARDGYSVTSLSRAHLAGOSILSGYSTY
 IYVATAPNMENVDLVGYSHPHYEOEVALGGIPLYQIIGWYRVNFGVIDERLHRN
 REYDRYRNLIAPAEDEGYRLAGPRPDQAREPPIHHAPOGCCNSSRTTGDTCN
 EETONLSTIYREYOSKVKROIYFSYOSEVDIYNIRDEL"
 BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN
 alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 us-09-528-682-1 x A04913 ..

Align seg 1/1 to: A04913 from: 1 to: 777

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
  |||
55 AATGGCGACAGATATATACCGTGTGACTGACCCAGATGAAATGAAATAA 104
  |||
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||
105 ACGTTCGGAGGTCTTATGCGGAGGCGCATATGAGTACTTCGATAG 154
  |||
34 LyrThrGluMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
155 GAACCAATGATGATATATCTTATGATTCACGGGAGGAGAACACAAACC 204
  |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
  |||
205 GCGTTTGTCAATATGATGACGAGATGTTTCCACTTCTTGTAGTTGAG 254
  |||
67 gSerAlaHisIleuAlaGlyInserIleLeuSerGlyTyrSerThrTyrT 84
  |||
255 AAGTCCCTACTAGCAGACAGTCTATATTCAGGATATCCACTTACT 304
  |||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
305 ATATATATGTTATAGCGACACACCAATATGTTTAAATGATATGATGTA 354
  |||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
  |||
355 TTAGGCGGTATACAGCCCTCACCCATATGACAGAGGTTTCTGCGTTAG 404
  |||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
  |||
405 TGGATFACCATATCTCTCAGATATATGATGATGATGATGATGATGATG 454
  |||
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||
455 TGATGATGAAAGCATTCATCGTAAACAGGAAATATAGAGCCGGAATTAC 504
  |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
  |||
505 AGAATTCGATATATAGCTCCGCGACAGAGATGTTACAGATATGACAGG 554
  |||
167 eProProAspHisGlnAlaTyrPargGluGluProTyrIleHisAlaP 184
  |||
555 CCCACCGGATCACCAAGCTTGAGAGAGAACCCGTGATCATCATGACAC 604
  |||
184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||
605 CACAAGGTTGTGGAATTCATCAAGAACATCACAGGTGATACCTTGTAAT 654
  |||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
  |||
655 GAGGAGACCCAGAAATCTGACACAAATATATCTCAGGAGAAATATCAATCAA 704
  |||
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||
705 AGTTAAGAGCGAGATATTTTCACTATACAGTTCAGAGGTTGACATATATA 754
  |||
234 sPargIleArgAspGluLeu 240
  |||
755 ACAGATTCGGGATGATTA 774
  |||

```

seq_name: gb_ba:ENSLTPA

seq_documentation_block:

LOCUS ENSLTPA 866 bp DNA linear BCT 26-Apr-1993
 DEFINITION Plasmid ENT P307 (enterotoxigenic E.coli) LTP subunit A gene.
 ACCESSION M15362
 VERSION M15362.1 GI:148337
 KEYWORDS LTP gene; heat-labile toxin.
 SOURCE Plasmid ENT (enterotoxigenic E.coli porcine isolate) DNA, clone p370.

ORGANISM Plasmid ENT
 plasmids.
 REFERENCE 1 (bases 1 to 866)
 AUTHORS Yamamoto, T., Gojohori, T. and Yokota, T.
 TITLE Evolutionary origin of pathogenic determinants in enterotoxigenic
 Escherichia coli and Vibrio cholerae O1
 JOURNAL J. Bacteriol. 169, 1352-1357 (1987)
 MEDLINE 87137303
 COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto
 (22-Apr-1987)

FEATURES

source location/Qualifiers
 1..866 /organism="Plasmid ENT"
 /plasmid="plasmid ENT"
 /db_xref="taxon:2463"
 sig_peptide 90..143
 CDS 90..866
 /note="Heat-labile toxin subunit A signal peptide"
 /note="Heat-labile toxin subunit A precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AAA24793.1"
 /db_xref="GI:148338"
 /translation="MKNITFIFPILASPLYANGRLYRASPDPDEIKRSGLMPLRG
 HNEYFDKQMNINIDHARGTGTGFEVRYDGYSTSLSSAHLAGOSILSGSTYI
 IYIVTAPNMFVNDVLGVSPHPYEDVSLAGIPIYQIIGWVRVNGVLDRLARN
 REYRDYRNLINIAPEDEGYRLAGFPDHOAWREBPWIHHAPOCGGNSRTITGPTCN
 EETONLSTIYREYQSKVKRQIFSDYQSEVDIYNRIDEL"
 mat_peptide 144..863
 /product="Heat-labile toxin subunit A"
 BASE COUNT 277 a 149 c 181 g 259 t
 ORIGIN unreported.

alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ENSLTPA ..

Align seg 1/1 to: ENSLTPA from: 1 to: 866

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
  |||
144 AATGGCGACAGATATATACCGTGTGACTGACCCAGATGAAATGAAATAA 193
  |||
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||
194 ACGTTCGGAGGTCTTATGCGGAGGCGCATATGAGTACTTCGATAGAG 243
  |||
34 LyrThrGluMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
244 GAACCTCAAATGATATTAATCTTATGATCAGCGGAGAGGAGACCAAAC 293
  |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
  |||
294 GCGTTTGTCAATATGATGACGAGATATGTTTCCACTTCTTGTAGTTGAG 343
  |||
67 gSerAlaHisIleuAlaGlyInserIleLeuSerGlyTyrSerThrTyrT 84
  |||
344 AAGTCTCAGTATGACGACAGAGTCTATATATGATGATATATCCACTTACT 393
  |||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
394 ATATATATGTTATAGCGACACCAATATGTTTAAATGTTAAATGATGTA 443
  |||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
  |||
444 TTAGGCGGTATACAGCCCTCACCCATATGACAGAGAGGTTTCTGCGTTAG 493
  |||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
  |||

```



```
|||||
494 TGAATACCATTTCTCAGATATATGATGCTATCCGTTAATTTGCTG 543
134 a11leaSPGLuArGLuEnHIsArGAsnArGLuTyArArGaSPArGTYrTYr 150
544 TGATTGATGACGATTACATCTGTAACAGCGAATATAGACCCGATTATAC 593
151 ArGAsnLeuSnHleAlaPrOAlaGLuASpGLyTyArArGLuAlaGLyPh 167
594 AGAATCTGAAATAGCTCCGCGAGAGATGTTACAGATTACACAGGTTT 643
167 ePrOPrOASpHISGLInAlaTrPARgLuGLuPrOTrPIleHIShIsAlaP 184
644 CCCACCGGATCACCAAGCTTGAGAGAAACCTGGATTCAATCATGAC 693
184 roGLInGLyCySGLyAsnSerSerArGThrlleThrGLyAspThrCyAsn 200
694 CACAGAGTTGTGAAATTCATCAAGAACATCACAGCTGATCTGTAAAT 743
201 GLUGluThrGLInAsnLeuSerThrlleTyArLeuArGLuTyArGLInSer 217
744 GAGGAGACCCAGAAATCTGACACAAATATATCTCAGGAAATATCAATCAA 793
217 sValTyArGLInlePheSerAspTYrGLInSerGLuValAspIleTyAr 234
794 AGTTAAGAGCGAGATATTTTCAGACTATCATGTCAGAGTTGACATATATA 843
234 snArGLleArGaSPGLuLeu 240
844 ACAGAAATTCGGGATGAATTA 863

seq_name: gb_ba:ENRLTPA

seq_documentation_block:
LOCUS      ENRLTPA                      898 bp    DNA     linear   BCT 26-APR-1993
DEFINITION Plasmid ENM-R PCG86 (enterotoxigenic E.coli) LTP subunit A gene.
ACCESSION  M15361
VERSION    M15361.1 GI:148333
KEYWORDS   LTP gene; heat-labile toxin.
SOURCE     Plasmid ENM-R (enterotoxigenic E.coli porcine isolate) DNA, clone
           PCG86.
ORGANISM   Plasmid ENM-R
REFERENCE  1 (bases 1 to 898)
AUTHORS    Yamamoto,T., Gojobori,T. and Yokota,T.
TITLE      Evolutionary origin of pathogenic determinants in enterotoxigenic
           Escherichia coli and Vibrio cholerae O1
JOURNAL    J. Bacteriol. 169, 1352-1357 (1987)
MEDLINE    87137303
COMMENT    Clean copy of sequence [1] kindly provided by T.Yamamoto
           (22-APR-1987).

FEATURES             Location/Qualifiers
     source           1..898
                     /organism="Plasmid ENM-R"
                     /plasmid="Plasmid ENM-R"
                     /db_xref="taxon:2462"
     sig_peptide      122..175
                     /note="Heat labile toxin subunit A signal peptide"
     CDS              122..898
                     /note="Heat labile toxin subunit A precursor"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA24791.1"
                     /db_xref="GI:148334"
                     /translation="WKNTFFIFLILASPLYANGDRLYRADSPPEDEIKRSGGLMPRG
                     HNEFYDRTOMNINLYDHARGTGTGFRYDGDYVTSLSAHLACGSIISGYSYTY
                     IYVATAPNMFVNVDLVGVSPHPEYGEVDAAGVPSQISGAWRVNFGVIDERLARN
                     REYRDYRYRNINIAPEADGYRPLAGFPDHOAMREPMWIIHHPQGCGNSRRTTGTCTN
                     EETQNLSTIYLREYQSKVRQIFSDYQSEVDIYNKIREL"
     mat_peptide      176..895
                     /product="Heat labile toxin subunit A"

BASE COUNT      289 a      155 c      186 g      268 t
ORIGIN           Unreported.
```

```
alignment_scores:
Quality: 240.00      Length: 240
Ratio: 1.000         Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ENRLTPA  ..

Align seq 1/1 to: ENRLTPA from: 1 to: 898

1  AsnGLyAspArGLuTyArArGLaAspSerArProPrOAspGLuIleTy 17
|||||
176  AATGGCGACAGATTATACCGTCTGACTTACAGCCCAATGAAATTA 225
17  sArGSerGLyGLyLeuMetPrOArGLyHIsAsnGLuTyArPheAspArG 34
226  ACGTTCGGAGGTCCTTATGCCAGAGGCAATATGACTACTTCGATTAG 275
34  LyThrGLInMetAsnHleAsnLeuTyArAspHIsAlaArGLyThrGLIn 50
276  GAACCTCAATGAATATTAACTTTATGATCAGCGAGAGAACCAACACC 325
51  GLyPheValArGLyTyArAspAspGLyTyArValSerThrSerLeuSer 67
326  GCGTTTCTCAGATATGATGACGAGATATGTTCCCTCTCTTACTTGAG 375
67  gSerAlaHIsLeuValArGLyGLInSerIleLeuSerGLyTyArSerThr 84
376  AAGGCTCACTTACAGAGACAGCTATATATGAGATATTCACCTTACT 425
84  yTrIleTyValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
426  ATATATATGTTATTAGCGACACCAAAATATGTTTAAATGTTAATGATGA 475
101  LeuGLyValTySerProHIsProTyrGLInGLuValSerAlaLeuGL 117
476  TTAGGCTATACAGCCCTCACCCATATGAACAGAGAGTTTCGGTTAGG 525
117  yGLyIleProTyrSerGLInIleTyrGLyTyArTyArValAsnPheGLy 134
526  TGAATATCAATATCTCAGATATGATGATGATGCTATCGTTAATTTGGTG 575
134  a11leaSPGLuArGLuEnHIsArGAsnArGLuTyArArGaSPArGTYrTYr 150
576  TGATTGATGAACGATTACATCTGTAACAGGAAATATAGAACCCGATTATAC 625
151  ArGAsnLeuSnHleAlaPrOAlaGLuASpGLyTyArArGLuAlaGLyPh 167
626  AGAATCTGAAATAGCTCCGCGAGAGATGTTTACAGATTACAGAGTTT 675
167  ePrOPrOASpHISGLInAlaTrPARgLuGLuPrOTrPIleHIShIsAlaP 184
676  CCCACCGGATCACCAAGCTTGAGAGAAACCTGGATTCAATCATGAC 725
184  roGLInGLyCySGLyAsnSerSerArGThrlleThrGLyAspThrCyAsn 200
726  CACAGAGTTGTGAAATTCATCAAGAACAAATCACAGCTGATCTGTAAAT 775
201  GLUGluThrGLInAsnLeuSerThrlleTyArLeuArGLuTyArGLInSer 217
776  GAGGAGACCCAGAAATCTGACACAAATATATCTCAGGAAATATCAATCAA 825
217  sValTyArGLInlePheSerAspTYrGLInSerGLuValAspIleTyAr 234
826  AGTTAAGAGCGAGATATTTTCAGACTATCATGTCAGAGTTGACATATATA 875
234  snArGLleArGaSPGLuLeu 240
876  ACAGAAATTCGGGATGAATTA 895

seq_name: gb_pat:E03421
```



```

seq_documentation_block:
  LOCUS      E03421              1148 bp    DNA          linear    PAT 29-SEP-1997
  DEFINITION  DNA encoding Lth of entero toxigenic Escherichia coli.
  ACCESSION  E03421
  VERSION    E03421.1 GI:2171637
  KEYWORDS   JP 1992079898-A/1.
  SOURCE     Escherichia coli.
  ORGANISM   Escherichia coli
              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
              Escherichia.
REFERENCE    1 (bases 1 to 1148)
  AUTHORS    Danbara,H. and Abe,A.
  TITLE      DNA AND RNA PROBE TO SIMULTANEOUSLY DETECT VIBRIO CHOLERAE AND
              ENTEROTOXIGENIC ESCHERICHIA COLI AND DETECTION OF VIBRIO CHOLERA
              AND ENTEROEXIGENIC ESCHERICHIA COLI USING THE SAME DNA AND RNA
              PROBE
  JOURNAL    Patent: JP 1992079898-A 1 13-MAR-1992;
              KITASATO INST.:THE
              OS      Escherichia coli
              PN      JP 1992079898-A/1
              PD      13-MAR-1992
              PF      23-JUL-1990 JP 1990194208
              PI      DANBARA HIROFUMI, ABE AKIO
              PC      C1201/68,C07H21/02,C07H21/04,C12N15/11,C12N15/31,C1201/04, PC
              C1201/10,
              PC      (C12N15/31,C12R1:19),(C1201/04,C12R1:63,C12R1:19); CC
              CC      strandedness: Double;
              CC      topology: Linear;
              CC      hypothetical: No;
              CC      anti-sense: No;
              CC      *source: clone-PKAD008;
              FH      Key      Location/Qualifiers
              FT      misc_feature      1..1148
                                   /note='DNA fragment encoding Lth of entero
                                   /toxicogenic
                                   /Escherichia coli'
              FT      sig_peptide      1..54
              FT      mat_peptide      55..774
                                   /product='subunit A of Lth'
              FT      CDS
              FT      sig_peptide      774..837
                                   /product='precursor subunit A of Lth'
              FT      mat_peptide      838..1145
                                   /product='subunit B of Lth'
              FT      CDS
                                   774..1148
              FT      CDS
                                   /product='precursor of subunit B of Lth'
              FEATURES
                source      1..1148
                                   /location/Qualifiers
                                   /organism='Escherichia coli'
                                   /db_xref='taxon:562'
  BASE COUNT      399 a      205 c      229 g      315 t
  ORIGIN
alignment_scores:
  Quality: 208.00      Length: 208
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-528-682-1 x E03421 ..
  Align seg 1/1 to: E03421 from: 1 to: 1148
  5 leuTyrAlaIaaspSerArgProProkaspGluIleuIysArgSerGlyGI 21
  67 TTATACCGTGTGACTAGACCCCGACAGATGAATAAATCAACCTCCGAGG 116
  21 YleuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetA 38
  117 TCTTATGCCAGAGGCGATTAATGAGTCTCGATAGAGGAAGAACTCAATGA 166

```

```

38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
167 ATATTATCTTTTATGATCAGCGAGAGAGACACAAACGGCTTTGTGCAGA 216
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisIle 71
217 TATGATGACGAGATATGTTTCCACTCTCTTAACTTGAAGTGCCTCAGCT 266
71 uAlaGlyGlnSerIleLeuSerGlyTyrSerThrTrpTyrIleTyrValI 88
267 AGCAGACAGCTATATATATGAGATATTCACCTTACTATATATATGTTA 316
88 leaIarTrAlaProAsnMetPheAsnValAsnAspValleuGlyValTyr 104
317 TAGCGACAGACCAATATGTTTATGTTATGATGATGATGAGCGATATAC 366
105 SerProHisProTyrGluGluGluValSerAlaLeuGlyGlyIleProty 121
367 AGCCCTCACCACCATATGAAACAGAGCGTTTCGCGTTAGGTGAAATACCA 416
121 rSerGlnIleTyrGlyTrpTyrTrpValAsnPheGlyValIleAspGluA 138
417 TTCTCAGATATATGATGATGATGATGATTAATTTGGTGTATGATGAAAC 466
138 rGleuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
467 GATTACATCGTAAACAGGAGATATAGACCGGTATTACAGAAATCTGAAAT 516
155 lIeAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProkaspH 171
517 ATACCTCGCGCAGAGAGATGTTACAGATTTAGCAGGTTTCCACCGGATCA 566
171 sGlnAlaTrpArgGluGluProTrpIleHisAlaProGluGlyCysG 188
567 CCAGACTTGAGACAGACACACCTGATTCATCATGCACCAAGGTTGTG 616
188 lYAsnSerSerArgThrIleThrGlyAspPheCysAsnGluTrpGln 204
617 GAATTCATCATGAAACAAATTAACAGGTGATCTGTATATGAGAGACCAG 666
205 AsnLeuSerThrIleTyrLeuArg 212
667 AATCTGAGCACAAATATATCTCAGG 690
seq_name: gb_ba:AB011677
seq_documentation_block:
  LOCUS      AB011677              1434 bp    DNA          linear    BCT 23-APR-1998
  DEFINITION  Escherichia coli genes for heat-labile enterotoxin A subunit and B
              subunit, complete cds.
  ACCESSION  AB011677
  VERSION    AB011677.1 GI:3062900
  KEYWORDS   heat-labile enterotoxin B subunit; Lth b subunit; heat-labile
              enterotoxin A subunit; Lth a subunit.
  SOURCE     Escherichia coli (strain:1032 (Enterotoxigenic)) DNA.
  ORGANISM   Escherichia coli
              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
              Escherichia.
REFERENCE    1 (bases 1 to 1434)
  AUTHORS    Komase,K.
  TITLE      Direct Submission
  JOURNAL    Submitted (02-MAR-1998) Katsuhiko Komase, The Kitasato Institute,
              Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
              Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-k@kitasato.or.jp,
              Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)
  2 (sites)
  Tamura,S., Asanuma,H., Tomita,T., Komase,K., Kawahara,K.,
  Danbara,H., Hattori,N., Watanabe,K., Suzuki,Y., Nagamine,T.,
  Aizawa,C., Oya,A. and Kurata,T.
  Escherichia coli heat-labile enterotoxin B subunits supplemented
  with a trace amount of the holotoxin as an adjuvant for nasal
  influenza vaccine

```


JOURNAL Vaccine 12 (12), 1083-1089 (1994)
 MEDLINE 95091056
 FEATURES
 source Location/Qualifiers
 1.1434
 /organism="Escherichia coli"
 /strain="1032 (Enterotoxigenic)"
 /db_xref="taxon:562"
 91.867
 /gene="Ltrh a subunit"
 91.867
 /gene="Ltrh a subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin A subunit"
 /protein_id="BAA25725.1"
 /db_xref="GI:3062901"
 /translation="MKNITFIFILLASPLYANGDKLYRADSRPDEIKRSGGLMPRG
 HNEYPDRGOMINILYDHAQGTQTFVRYDDGVSTSLRSALHLAGOSILSGSYTY
 IYVIATAPMFMVNDVLYGYSPPHYEQEYSAAGITPSQICMYRNVNPGVIDERLHRN
 REYRDYRNLIAPAEDEGRLAGPPPDQANREPWIIHAPQGCGDSSRTITTDTCN
 EETONLSTIYLKRYOSKVRQIFSDYQSEVDIYNRIDEL"
 864.1238
 /gene="Ltrh b subunit"
 864.1238
 /gene="Ltrh b subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin B subunit"
 /protein_id="BAA25726.1"
 /db_xref="GI:3062902"
 /translation="MNKVKCYVLFALASSICAGAPOSITELCSFRMTQITINDK
 ILSYESNAGKREMYITPKSGATFOVEPFGSOHIDSQKALIERMKDITRLITLTETK
 IDKLQVNNKRTNSIALISMEN"
 BASE COUNT 488 a 258 c 286 g 402 t
 ORIGIN
 Alignment_scores:
 Quality: 191.00 Length: 191
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 Alignment_block:
 US-09-528-682-1 x AB011677 ..
 Align seg 1/1 to: AB011677 from: 1 to: 1434
 5 LeuTYrAGAlaAspSerArgProProAspGluIleuYsArgSerGlyGI 21
 ||||||||||||||||||||||||||||||||||||||||||||||||
 157 TTATACGCTGCTGACTCTAGACCCCGCATGAATAAAGCTTCGGAGG 206
 21 YLeuMetProArgGlyHisAsnGluTYrPheAspArgGlyThrGlnMetA 38
 ||||||||||||||||||||||||||||||||||||||||||||||||
 207 TCTTATGCCAGAGCGCATATGAGTACTTCGATAGAGAACTCAATAGA 256
 38 snIleAsnLeuTYrAspHisAlaArgGlyThrGlnIleuPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 257 ATATTAATCTTATGATCACCAGAGAACACAAACGGCTTGTGCGA 306
 55 TYrAspAspGlyTYrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
 ||||||||||||||||||||||||||||||||||||||||||||||||
 307 TATGATGACGATATGTTTCCACTTCCTTACTTGTGAGAACTGCTCACTT 356
 71 AAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYrTleTYrValI 88
 ||||||||||||||||||||||||||||||||||||||||||||||||
 357 AGCAGGACAGCTATATTATCAGATATCCACTTACTATATATATGTTA 406
 88 IeAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTYr 104
 ||||||||||||||||||||||||||||||||||||||||||||||||
 407 TAGGACAGCACCAATATGTTTAAATGTTAAATGATTAAGCGCATAC 456
 105 SerProHisProTYrGluGlnGluValSerAlaLeuGlyGlyIleProTY 121
 ||||||||||||||||||||||||||||||||||||||||||||||||

457 AGCCCTCACCATATGAAACAGAGGTTTCGCTTAGGTGATACATA 506
 121 rSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyValIleAspGluA 138
 ||||||||||||||||||||||||||||||||||||||||||||||||
 507 TTTCAGATATATGATGATGCTGCTTAAATTTGCTAATTAATGATGAC 556
 138 rGluHisArgAsnArgGluTYrTYrArgAspArgTYrTYrArgAsnLeuAsn 154
 ||||||||||||||||||||||||||||||||||||||||||||||||
 557 GATTACATCGTACACGAGATATAGACCGGCTATTACAGAAATCTTGAT 606
 155 ILeAlaProAlaGluAspGlyTYrArgGluValAlaGlyPheProProAspH 171
 ||||||||||||||||||||||||||||||||||||||||||||||||
 607 ATACTCGGACAGAGATGTTTACAGATTACAGGTTTCCACCGGATCA 656
 171 sGlnAlaTPrArgGluGluProTPrIleHisAlaProGlnGlyCysG 188
 ||||||||||||||||||||||||||||||||||||||||||||||||
 657 CCAAGCTTGAGAGAGAACCCGTGATTCATCAGACACACAGGTTGTG 706
 188 LysAsnSerSerArgThrIleThr 195
 707 GAAATTCATCAAGAACCAATTACA 729
 seq_name: gb_ba:ECOTOXA

seq_documentation_block: 934 bp DNA linear BCT 26-APR-1993
 LOCUS ECOTOXA
 DEFINITION E.coli toxa gene encoding subunit A of heat-labile enterotoxin.
 ACCESSION K01995
 VERSION K01995.1 GI:148027
 KEYWORDS enterotoxin; heat-labile enterotoxin; toxa gene.
 SOURCE E.coli H10407 (serotype 078:H11) plasmid DNA, clone pUY27.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 934)
 YAMAMOTO,T., Tamura,T. and Yokota,T.
 Primary structure of heat-labile enterotoxin produced by
 Escherichia coli pathogenic for humans
 J. Biol. Chem. 259, 5037-5044 (1984)
 MEDLINE 84185610
 COMMENT This heat-labile enterotoxin is pathogenic for humans (Ltrh).
 Compared in [1] with LTP A and CT A (LTP-pathogenic for piglets,
 CT-cholera toxin).
 CT-cholera toxin)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 FEATURES
 source Location/Qualifiers
 1.934
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 158.934
 /note="heat-labile enterotoxin A prepeptide"
 /codon_start=1
 /transl_table=11
 /protein_id="AA24685.1"
 /db_xref="GI:148028"
 /translation="MKNITFIFILLASPLYANGDKLYRADSRPDEIKRSGGLMPRG
 HNEYPDRGOMINILYDHAQGTQTFVRYDDGVSTSLRSALHLAGOSILSGSYTY
 IYVIATAPMFMVNDVLYGYSPPHYEQEYSAAGITPSQICMYRNVNPGVIDERLHRN
 REYRDYRNLIAPAEDEGRLAGPPPDQANREPWIIHAPQGCGDSSRTITTDTCN
 EETONLSTIYLKRYOSKVRQIFSDYQSEVDIYNRIDEL"
 158.211
 /note="heat-labile enterotoxin A signal peptide"
 212.787
 /product="heat-labile enterotoxin A1"
 794.931
 /product="heat-labile enterotoxin A2"
 165 c 186 g 285 t
 BASE COUNT 298 a 165 c 186 g 285 t
 ORIGIN 46 bp upstream of FokI recognition pattern.

alignment_scores:
 Quality: 184.00 Length: 184
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOTOXA

Align seg 1/1 to: ECOTOXA from: 1 to: 934

```

5 LeuTYRtRgAlaAspSerArgProProAspGluIleLysArgSerGlyG1 21
|||||
224 TTATACCGTGTGACTAGACCCCGAGATGAATAAAGCGTCCGAGG 273
21 YLeuMetProArgGlyHisAsnGluTYRPhesArgGlyThrGlnMetA 38
|||||
274 TCTTATGCCAGAGCGCATATAGTACTTCGATAGAGAACTCAATGA 323
38 snlleAsnLeuTYRAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
324 ATATTATCTTTATGATCAGCCGAGAGAACACAAACCGCTTGTCAGA 373
55 TYRAspArgGlyTYRValSerThrSerLeuSerLeuArgSerAlaHisLe 71
|||||
374 TAGATGACGGATATGTTCCACTTCCTTAGTTGAGAAAGTGCCTACTT 423
71 uAlaGlyGlnSerIleLeuSerGlyTYRSerThrTYRTrIleTYRValI 88
|||||
424 AGCAGACAGCTATATTATCAGATATTCACCTTACTATATATGTTA 473
88 leaIatHraIaProAsnMetPheAsnValAsnAspValLeuGlyValTYR 104
|||||
474 TAGGACAGACACCAATATGTTTAAATGTTAATGATGATATGAGCGTATAC 523
105 SerProHisProTYRgInGluGlnValSerAlaLeuGlyIleProTY 121
|||||
524 AGCCCTACCCATATGAAACAGAGAGTTTCGCGTAGGTGGAATACCATTA 573
121 rSerGlnIleTYRgIleTYRTrIleTYRArgValAsnPheGlyValIleAspGluA 138
|||||
574 TTCACAGATATATGAGATGATATCGTTAAATTTGGTGATGATTCATGAC 623
138 rGleuHisAlaArgAsnArgGluTYRArgAspArgTYRTrIleTYRAsnLeuAsn 154
|||||
624 GATTACATCGTTAACAGGAAATATAGACCCGCTATTACAGAAATCTGAAT 673
155 lleaIaProAlaGluAspGlyTYRArgLeuAlaGlyPheProProAspH 171
|||||
674 ATACCTCGGAGAGAGATGTTACAGATTAGCAGGTTCCACCGGATCA 723
171 sGlnAlaTrpArgGluGluProTrIleHisHisAlaProGlnGlyCysG 188
|||||
724 CCAAGCTTGGAGAGAGAACCCGTGATTCAATCATGCAACCAAGGTGTG 773
188 ly 188
774 GA 775

seq_name: gb_ba:S60731
seq_documentation_block:
LOCUS S60731 1275 bp DNA linear BCT 29-JAN-2002
DEFINITION heat-labile enterotoxin A subunit, heat-labile enterotoxin B
subunit [Escherichia coli, 21d, Genomic, 1275 nt].
ACCESSION S60731
VERSION S60731.1 GI:408994
KEYWORDS
SOURCE Escherichia coli 21d.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 1275)
AUTHORS Inoue,T., Tsuji,T., Koto,M., Iinamura,S. and Miyama,A.
TITLE Amino acid sequence of heat-labile enterotoxin from chicken
enterotoxigenic Escherichia coli is identical to that of human
strain H 10407
JOURNAL FEMS Microbiol. Lett. 108 (2), 157-161 (1993)
MEDLINE 93252225

```

PUBMED

8486242

REMARK

Genbank staff at the National Library of Medicine created this entry [NCBI gisbseq 131491] from the original journal article. This sequence comes from Fig. 2.

FEATURES

Location/Qualifiers

1..1275

/organism="Escherichia coli"

/db_xref="taxon:562"

68..844

/note="This sequence comes from Fig. 2; conceptual translation presented here differs from translation in publication; LTC A subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin A subunit"

/protein_id="AAC60440.2"

/db_xref="GI:18390349"

/translation="MKNTFEIFILLASPIYANGDKLYRADRPDEIKRSGGLMPRG HNEFYDRGTOMKINILYDARGTQTFVYRDDGYTSLSLSAHLAGSILSGSYX IYVATAPRMENVNDLVYSPHPTFEDVSALGILYISQITGMTRVNFVYIDRLHRN REYRDYRYRNLIAPAEDEGYRLAGEPPDHQAMREBPWLIHAPQGGDESSRTTGDTCN EETQWLSYILKYSKYVRQIFSDYQSEVDIYNIRNEL"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCEYRNTOIYTLNDK ILSYTESNAGKREMYITPKSGATQVEYPSQHIDSQKAIERKDKLRLTYLLETK IDKLGVNKKTFPNSIAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLCAGAPOSITELCE

seq_name: gb_pat:A04915			
seq_documentation_block:			
LOCUS	A04915	777 bp	DNA
			linear
			PAT 14-JUL-1993

seq_documentation_block:	777 bp	DNA linear	PAT 14-JUL-1993
LOCUS	A04915		
DEFINITION	Artificial sequence	LTA gene for labile toxin A.	
ACCESSION	A04915		
VERSION	A04915.1	GI:412519	
KEYWORDS	labile toxin A.		
SOURCE	synthetic construct.		
ORGANISM	synthetic construct.		
REFERENCE	artificial sequence.		
AUTHORS	1 (bases 1 to 777)		
TITLE	Hayes,M.V., Hatford,S. and Ross,G.W.		
JOURNAL	Microbiological process		
	Patent: EP 0145486-A 7 19-JUN-1985;		

FEATURES	Location/Qualifiers
source	1. .777

CDS

```

BASE COUNT      255 a      135 c      164 g      222 t      1 others
ORIGIN
//cocon_table=1
//transl_table=1
/product="labile toxin A"
/protein_id="CAA00403.1"
/db_xref="GI:412520"
/translation="MKNIITFFILASPLYANGDRLYRADSRPPEDEIKRSGLMPCG
HNEFDRGTGMNINLYDHARGTQGFVARYDGYXTSLSLASALACGSLISGTYSTY
IYVATAPNMFNVNDVIGVYSPHYEDEVSLGIPISQITGVRVNGCVTDEELHRH
REYRDQRYRRLNTAPADGGRVLAGFPFDPDHQAMREPEPNIHAPQCGGNSRITTDGTC
EFTQNLSTIYLRQTSVKRQFESDVQSEVDIINRDEL"

```

```

alignment_scores:
  Quality: 179.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 179
  Gaps: 0
  Percent Identity: 100.000

```

alignment_block:
US-09-528-682-1 x A04915

Align seg 1/1 to: A04915 from: 1 to: 777

62 ~~Thrsr~~msr~~se~~rla~~rs~~rs~~er~~la~~h~~is~~re~~la~~l~~actylins~~er~~leu~~se~~ 78
238 ~~ACTCTCTAGTTAGTTGAGAGGCTCACTTACAGGACAGCTCATATTATC~~ 287
78 rcl~~y~~ty~~r~~ser~~th~~ty~~r~~ty~~r~~le~~y~~r~~v~~all~~e~~al~~th~~al~~a~~pro~~as~~me~~lp~~ 95
288 ~~AGCAATTCACACTACTCATATATATGTTATGGCAGCAGCACCAATATGT~~ 337

95 heasnvalasnaspyaleucglylalytserprohisprotyrarglun 1111
338 tyyaaagttaaagatctatattagccgattacacccctccacccataatgaacag 3877
112 giuvalasratialeuglylileprotyserscintletrygltyrpy 1287
388 caggtgttccgcttagtgggaattaccattctcagattatattgatgtga 4373
128 taryvalasnbheglylalleaspgluatrgleuhsirgasmarglut 1457
438 tgcgtttatttttggcttgattgataagaaacgattacatcgctaacggcaat 4873
145 yfarasparctytcttyrargasneusnillealaprolagiaspgly 1617
488 atnagagaccggattatcagaatctgaaattatagctccggcagagatgct 5373
162 tyratrgleualaglypheproproasphisinalatrrpargluglut 1787
538 ttcacatttagcaggtttcccaaccgattccaccaccttgagagaaagAAC 5877
178 crrprrlenshsialaproginglycyasglyasnsersertrghrllet 1957
588 ctggattatctatctgcaccaccaaaggttgaggaaattcatcaaaAACATCA 6373
195 hrGlyAsPthrcysasngluclutrhglnasleuSerthrIleTyrLeu 2117
638 CAGGTGAATACCTGTATATGAGAGAACCCAGAAATCTGAGACCAATATATCTC 6877
212 ArGgLUtyrGlnSerIleValIysArgIlnIlePheSerAspTyrGlnse 2287
688 AGGGAATATCAATCAAAAGTTTACAGGCAGAGATATTTTCAAGCTATCAATC 7373
228 rgluValAspIleTyrAsnArgIleArgAspGluLeu 240
|||||
738 AGAGCTTACACATATTTAACAAATTCGGGATGAAATTTA 774

seq_name: gqb_ba:EWDLTA

seq_documentation_block:

LOCUS	723 bp	DNA	linear	BCT 23-APR-1996
DEFINITION	EMD299 (from E. coli)	heat-labile enterotoxin subunit A		
	plasmid			
	(LT-A) gene, 3' end.			

ACCESSION	M57244 J05702
VERSION	M57244.1 GI:531190
KEYWORDS	enterotoxin; heat-labile enterotoxin
SOURCE	Plasmid EMD 299 DNA.
ORGANISM	Plasmid EMD 299

REFERENCE
1 (bases 1 to 723)

AUTHORS Tsujii, T., Inoue, T., Miyama, A., Okamoto, K., Honda, T. and Miyetani, T.
TITLE A single amino acid substitution in the A subunit of *Escherichia coli* enterotoxin results in a loss of its toxic activity
JOURNAL J. Biol. Chem. 265 (36), 22520-22525 (1990)

COMMENT	On Aug 20, 1994 this sequence version replaced g1:148531
FEATURES	Location/Qualifiers

FEATURES

Source

Source

gene
CDS

CDS

```

1. 723
2. 723
3. 723
4. 723
5. 723
6. 723
7. 723
8. 723
9. 723
10. 723
11. 723
12. 723
13. 723
14. 723
15. 723
16. 723
17. 723
18. 723
19. 723
20. 723
21. 723
22. 723
23. 723
24. 723
25. 723
26. 723
27. 723
28. 723
29. 723
30. 723
31. 723
32. 723
33. 723
34. 723
35. 723
36. 723
37. 723
38. 723
39. 723
40. 723
41. 723
42. 723
43. 723
44. 723
45. 723
46. 723
47. 723
48. 723
49. 723
50. 723
51. 723
52. 723
53. 723
54. 723
55. 723
56. 723
57. 723
58. 723
59. 723
60. 723
61. 723
62. 723
63. 723
64. 723
65. 723
66. 723
67. 723
68. 723
69. 723
70. 723
71. 723
72. 723
73. 723
74. 723
75. 723
76. 723
77. 723
78. 723
79. 723
80. 723
81. 723
82. 723
83. 723
84. 723
85. 723
86. 723
87. 723
88. 723
89. 723
90. 723
91. 723
92. 723
93. 723
94. 723
95. 723
96. 723
97. 723
98. 723
99. 723
100. 723
101. 723
102. 723
103. 723
104. 723
105. 723
106. 723
107. 723
108. 723
109. 723
110. 723
111. 723
112. 723
113. 723
114. 723
115. 723
116. 723
117. 723
118. 723
119. 723
120. 723
121. 723
122. 723
123. 723
124. 723
125. 723
126. 723
127. 723
128. 723
129. 723
130. 723
131. 723
132. 723
133. 723
134. 723
135. 723
136. 723
137. 723
138. 723
139. 723
140. 723
141. 723
142. 723
143. 723
144. 723
145. 723
146. 723
147. 723
148. 723
149. 723
150. 723
151. 723
152. 723
153. 723
154. 723
155. 723
156. 723
157. 723
158. 723
159. 723
160. 723
161. 723
162. 723
163. 723
164. 723
165. 723
166. 723
167. 723
168. 723
169. 723
170. 723
171. 723
172. 723
173. 723
174. 723
175. 723
176. 723
177. 723
178. 723
179. 723
180. 723
181. 723
182. 723
183. 723
184. 723
185. 723
186. 723
187. 723
188. 723
189. 723
190. 723
191. 723
192. 723
193. 723
194. 723
195. 723
196. 723
197. 723
198. 723
199. 723
200. 723
201. 723
202. 723
203. 723
204. 723
205. 723
206. 723
207. 723
208. 723
209. 723
210. 723
211. 723
212. 723
213. 723
214. 723
215. 723
216. 723
217. 723
218. 723
219. 723
220. 723
221. 723
222. 723
223. 723
224. 723
225. 723
226. 723
227. 723
228. 723
229. 723
230. 723
231. 723
232. 723
233. 723
234. 723
235. 723
236. 723
237. 723
238. 723
239. 723
240. 723
241. 723
242. 723
243. 723
244. 723
245. 723
246. 723
247. 723
248. 723
249. 723
250. 723
251. 723
252. 723
253. 723
254. 723
255. 723
256. 723
257. 723
258. 723
259. 723
260. 723
261. 723
262. 723
263. 723
264. 723
265. 723
266. 723
267. 723
268. 723
269. 723
270. 723
271. 723
272. 723
273. 723
274. 723
275. 723
276. 723
277. 723
278. 723
279. 723
280. 723
281. 723
282. 723
283. 723
284. 723
285. 723
286. 723
287. 723
288. 723
289. 723
290. 723
291. 723
292. 723
293. 723
294. 723
295. 723
296. 723
297. 723
298. 723
299. 723
300. 723
301. 723
302. 723
303. 723
304. 723
305. 723
306. 723
307. 723
308. 723
309. 723
310. 723
311. 723
312. 723
313. 723
314. 723
315. 723
316. 723
317. 723
318. 723
319. 723
320. 723
321. 723
322. 723
323. 723
324. 723
325. 723
326. 723
327. 723
328. 723
329. 723
330. 723
331. 723
332. 723
333. 723
334. 723
335. 723
336. 723
337. 723
338. 723
339. 723
340. 723
341. 723
342. 723
343. 723
344. 723
345. 723
346. 723
347. 723
348. 723
349. 723
350. 723
351. 723
352. 723
353. 723
354. 723
355. 723
356. 723
357. 723
358. 723
359. 723
360. 723
361. 723
362. 723
363. 723
364. 723
365. 723
366. 723
367. 723
368. 723
369. 723
370. 723
371. 723
372. 723
373. 723
374. 723
375. 723
376. 723
377. 723
378. 723
379. 723
380. 723
381. 723
382. 723
383. 723
384. 723
385. 723
386. 723
387. 723
388. 723
389. 723
390. 723
391. 723
392. 723
393. 723
394. 723
395. 723
396. 723
397. 723
398. 723
399. 723
400. 723
401. 723
402. 723
403. 723
404. 723
405. 723
406. 723
407. 723
408. 723
409. 723
410. 723
411. 723
412. 723
413. 723
414. 723
415. 723
416. 723
417. 723
418. 723
419. 723
420. 723
421. 723
422. 723
423. 723
424. 723
425. 723
426. 723
427. 723
428. 723
429. 723
430. 723
431. 723
432. 723
433. 723
434. 723
435. 723
436. 723
437. 723
438. 723
439. 723
440. 723
441. 723
442. 723
443. 723
444. 723
445. 723
446. 723
447. 723
448. 723
449. 723
450. 723
451. 723
452. 723
453. 723
454. 723
455. 723
456. 723
457. 723
458. 723
459. 723
460. 723
461. 723
462. 723
463. 723
464. 723
465. 723
466. 723
467. 72
```


old_sequence
459
/gene="JT-A"
/note="a"
/citation="[1]
/replace="a"

BASE COUNT 238 a 128 c 160 g 197 t
ORIGIN

alignment_scores:
Quality: 111.00 Length: 111
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x EMDLTA ..

Align seg 1/1 to: EMDLTA from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
1 AATGGGACGAGATTATACGCTGCTAGACCCAGATGAAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 ACCTTCCGAGGCTCTATGCGCCAGAGGCGCATATGACTACTTGCATRGAG 100
34 LyrThrGlnMetAsnLeuAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GAACCTCAATGAGATTATATCTTATGATCAGCCGAGAGACAAAC 150
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeuArg 67
151 GGCTTTGTCGATATGATGACGAGATAGTTTCCACTTCTCTAGTTGAG 200
67 gSerAlaHisIleGlyAlaGlyIleSerIleLeuSerGlyTyrSerThrTyrT 84
201 AATGCTCATCTTATACGACAGCTCTATATATCAGATATCCACTTACT 250
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATATAGCGACAGACCAAAATATGTTTATATGTTAATGATGTA 300
101 LeuGlyValTyrSerProHisProTyrGluGln 111
301 TTAGGCTATACGCCCTCACCATTATGAAACAG 333

seq_name: gb_pat:E00613

seq documentation_block:

LOCUS E00613 777 bp DNA linear PAT 29-SEP-1997
DEFINITION Dna encoding A component of heat labile toxin, LTA.
ACCESSION E00613
VERSION E00613.1 GI:2168892
KEYWORDS JP 1986005097-A/1.
SOURCE Escherichia coli.
ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE 1 (bases 1 to 777)

AUTHORS Maikuru,B.H., Suteilbun,H. and Goodon,U.R.
TITLE MICROBIAL METHOD
JOURNAL Patent: JP 1986005097-A 1 10-JAN-1986;

COMMENT GLAXO GROUP LTD
OS Escherichia coli
PN JP 1986005097-A/1
PD 10-JAN-1986
PF 12-DEC-1984 JP 1984262645
PR 12-DEC-1983 GB 83 8333131
PI MAIKURU BIKUTAA HEIZU, SUTEILIBUN HAFUODOO,

PI GOODON URIAMU ROSU
PC C07K15/04,A61K39/108,C12N1/20,C12N15/00//C12P21/00,(C12N1/20,
PC C12R1:19),
PC (C12N15/00,C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH CDS 1..777
FT /product='A component of heat labile FT
FT mat_peptide 1..774
FT /product='A component of heat labile FT
FT toxin,LTA' /product='A component of heat labile FT
FT mutation toxin,LTA' /product='A component of heat labile FT
FT Location/Qualifiers
FT 1..777
FT /organism='Escherichia coli'
FT /db_xref='taxon:562'

alignment_scores:
Quality: 97.00 Length: 97
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x E00613 ..

Align seg 1/1 to: E00613 from: 1 to: 777

75 SerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAl 91
277 TCATATATACAGATATCCACTTACTATATATATGTTATATAGCGAC 326
91 aProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerProHisP 108
327 ACCAAATATGTTTAAATGTTAATGATGATATGATGATGATGATGATG 376
108 rOTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnIle 124
377 CATATGACAGAGGTTTTCGCTTAGTGATGATACATATCTCAGATA 426
125 TyrGlyTyrPtyrArgValAsnPheGlyValIleAspGluArgLeuHis 141
427 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 476
141 gAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaPro 158
477 TAACAGGAGATATAGAGACCGGATATACAGAAATCGAATATAGCTCCG 526
158 IaGluAspGlyTyrArgLeuAlaGlyPheProProAspHis 171
527 CAGAGATGTTTACAGATTAGCAGGTTTCCACCGGATCAG 567

seq_name: gb_pat:A16419

seq documentation_block:

LOCUS A16419 711 bp DNA linear PAT 06-JUL-1995
DEFINITION A subunit of a labile toxin L71.
ACCESSION A16419
VERSION A16419.1 GI:641008
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE 1 (bases 1 to 711)

AUTHORS
TITLE IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN
JOURNAL LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES
PATENT: WO 9313202-A 42 08-JUL-1993;
FEATURES Location/Qualifiers
source 1..711
/organism="Escherichia coli"
/db_xref="taxon:562"
gene 1..711
/gene="LT1 subunit A"
CDS <1..711
/gene="LT1 subunit A"
/codon_start=1
/transl_table=11
/protein_id="CAA01280.1"
/db_xref="GI:641009"
/translation="NGDRLYRADSRPPEIKRFRSLMRPNENYEDRGTOINILYDHA
RGTOGEVRYDDGYVSTSLSAHLAQYLISGLTYIYIAMPFNVDIYSYP
HPYEEVSAIGIPYSOITGYWYRVNFGYIDELRHNRREYRNLMINAPEDGYRL
AGFPDHOAMREEPWIIHAPQCGDSSRTITGDCNETQNSTIYLREYOSKVKROI
FSDYQSEVDIYNRIDEL"
BASE COUNT 234 a 123 c 156 g 198 t
ORIGIN
alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x A16419 ..
Align seg 1/1 to: A16419 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIle 119
|||||
295 GTATACAGCCCTCACCATATGAACAGAGGTTTCGCTTAGGTGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyValIle 136
|||||
345 ACCATATTCACATATATGATGATGCTGTTAATTTTGGTGTGATG 394
|||||
136 spGUATgLeuHISArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACACATTACATCGTACACAGGAATATAGACCGGTATTCAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
445 CTGAATATACCTCCGCGCAGAGATGTTACAGATTACAGGTTCCACAC 494
|||||
169 OAEPHISGLAlaIleTyrArgGluGluProTyrPheHisAlaProGln 186
|||||
495 GGATTCACCAAGCTTGGAGAGAGAACCTGATTCATCATGACACAG 544
|||||
186 lYcysGly 188
|||||
545 GTTGTGCA 552
seq_name: gb_pat:AR118597
seq_documentation_block:
LOCUS AR118597 711 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6149919.
ACCESSION AR118597
VERSION AR118597.1 GI:14100507
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 711)
AUTHORS Domenighini, M., Rappuoli, R., Pizza, M. and Hol, W.
TITLE Immunogenic detoxified mutants of cholera toxin and of the toxin

JOURNAL LT, their preparation and their use for the preparation of vaccines
PATENT: US 6149919-A 5 21-NOV-2000;
FEATURES Location/Qualifiers
source 1..711
/organism="unknown"
BASE COUNT 234 a 123 c 156 g 198 t
ORIGIN
alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AR118597 ..
Align seg 1/1 to: AR118597 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIle 119
|||||
295 GTATACAGCCCTCACCATATGAACAGAGGTTTCGCTTAGGTGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyValIle 136
|||||
345 ACCATATTCACATATATGATGATGCTGTTAATTTTGGTGTGATG 394
|||||
136 spGUATgLeuHISArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACACATTACATCGTACACAGGAATATAGACCGGTATTCAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
445 CTGAATATACCTCCGCGCAGAGATGTTACAGATTACAGGTTCCACAC 494
|||||
169 OAEPHISGLAlaIleTyrArgGluGluProTyrPheHisAlaProGln 186
|||||
495 GGATTCACCAAGCTTGGAGAGAGAACCTGATTCATCATGACACAG 544
|||||
186 lYcysGly 188
|||||
545 GTTGTGCA 552
seq_name: gb_ba:ECELTA
seq_documentation_block:
LOCUS ECELTA 868 bp DNA linear BCT 12-SEP-1993
DEFINITION E. coli gene eltA encoding the A subunit of the heat-labile
enterotoxin (LT).
ACCESSION V00275.1 GI:11339
VERSION V00275.1 GI:11339
KEYWORDS enterotoxin; signal peptide.
SOURCE Escherichia coli.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 868)
AUTHORS Spicer, E.K. and Noble, J.A.
TITLE Escherichia coli heat-labile enterotoxin. Nucleotide sequence of
the A subunit gene
JOURNAL J. Biol. Chem. 257 (10), 5716-5721 (1982)
MEDLINE 82167425
FEATURES Location/Qualifiers
source 1..868
/organism="Escherichia coli"
/db_xref="taxon:562"
mrna 56..>868
/product="messenger RNA"
sig_peptide 90..143
CDS 90..854
/codon_start=1
/transl_table=11
/product="enterotoxin subunit A"


```
/protein_id="CAA23532.1"
/db_xref="GI:41340"
/translation="MKNTFFPILASLYANGRLYRADSRPPEDEIKRFRSLPRG
NEFDGTOININLDHARGTQTFVRIDGYVTSLSRSHLAGQYLLSGYSLLTY
IVIANFNVNDIVSYSPHPYQEVSAAGIYVSOIYGMVRYNFGVIDRLHREYR
DYYRNLNAPAEDEGYRLAGPPDHQAMREPEWIIHAPGCGDSSRTLTGDCNEETQ
NLSTYLRFYQSKVKRQIOPSDYQSEVDIYNRIRDEL"
BASE COUNT      284 a      142 c      179 g      263 t
ORIGIN

alignment_scores:
  Quality:      86.00      Length:      86
  Ratio:        1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x ECElTA ..

Align seg 1/1 to: ECElTA from: 1 to: 868

103 ValtyrSerProHisProtyrGluGlnGluValSerAlaLeuGlyIle 119
|||||
438 GTATACAGCCCTCACCATATGAAACAGAGAGGTTCTGCGTAGCGTGAA 487
|||||
119 eProtyrSerGlnIleTyrglyTrpTyrglyValAsnPhedGlyValIle 136
|||||
488 ACCATATTCAGATATATGATGATGATCGTTAAATTTGGTGTGATG 537
|||||
136 spGluAlaGluLeuHisArgAsnArgIuTyrglyArgAspArgTyrglyArgAsn 152
|||||
538 ATGAACGATTCATCATCGTAAACAGGAATATAGACACCGGTATTACAGAAAT 587
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrglyArgLeuAlaGlyPheProPr 169
|||||
588 CTGAATATAGCTCCGCCAGAGATGTTACAGATTAGCAGGTTCCACACC 637
|||||
169 oAspHisGlnAlaIleAlaTrpArgGluGluProTrpIleHisAlaProGln 186
|||||
638 GCATCACCACCAACGTTGCAGAGAGAACAACCTGATTCATCATCACCACAG 687
|||||
186 lYcysGly 188
|||||
688 GTTGTGGA 695

seq_name: gb_ba:ECLT87

seq_documentation_block:
  LOCUS      ECLT87              595 bp      DNA      linear      BCT 28-OCT-1996
  DEFINITION  E.coli LT87 gene for heat-labile enterotoxin.
  ACCESSION  X83966
  VERSION    X83966.1 GI:1648865
  KEYWORDS   heat-labile enterotoxin; LT 87 gene.
  SOURCE     Escherichia coli.
  ORGANISM   Escherichia coli
              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
              Escherichia.
  REFERENCE  1 (bases 1 to 595)
  AUTHORS    Germani,Y. and Desperrier,J.M.
  TITLE      Nucleotide sequence variations in genes encoding heat labile
              enterotoxins of Escherichia coli isolated in South Pacific
              unpublished
  JOURNAL    2 (bases 1 to 595)
  REFERENCE  Germani,Y.
  AUTHORS    Direct Submission
  TITLE      Submitted (17-JAN-1995). Y. Germani, Institut Pasteur, 28 rue du Dr
              Roux, F- 75734 Paris Cedex 15, FRANCE
  COMMENT    Related sequence: M15361-3 (Yakamoto).
  FEATURES   Location/Qualifiers
              1..595
              /organism="Escherichia coli"
              /strain="ETEC LT 87"
```

```
/db_xref="taxon:562"
gene
  205..579
  /gene="LT 87"
CDS
  205..579
  /gene="LT 87"
  /codon_start=1
  /transl_table=11
  /product="heat-labile enterotoxin"
  /protein_id="CAA58800.1"
  /db_xref="GI:1648866"
  /db_xref="SWISS-PROT:P13811"
  /translation="MKVKFYLFETALLSSLCAGAPQSITELCEYHNTIYIINDK
  ILSYTESMAGKREMYITFEKSGATFOVEVPSGHIIDSOKAIERRKDTLRITLYLETK
  IDKLCVMNNKTPNSIAISMEN"
BASE COUNT      226 a      105 c      113 g      151 t
ORIGIN

alignment_scores:
  Quality:      34.00      Length:      34
  Ratio:        1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x ECLT87 ..

Align seg 1/1 to: ECLT87 from: 1 to: 595

173 ALATPArgGluGluProTrpIleHisAlaProGlnGlycysGlyAs 189
|||||
3 GCTTGGAGAGAGAACCCCTGATTCATCATGACACCAAGGTTGTGAAA 52
|||||
189 nSerSerArgThrIleHrglyAspThrCysAsnGluGluThrGlnAsn 206
|||||
53 TTGATCAGAGACATTTACAGTCACTTGTAAATGAGAGACCCAGAAATC 102
|||||
206 eu 206
||
103 TG 104

seq_name: gb_pat:A06082

seq_documentation_block:
  LOCUS      A06082              791 bp      DNA      linear      PAT 25-MAY-1993
  DEFINITION  PEK620 HindIII-fragment.
  ACCESSION  A06082
  VERSION    A06082.1 GI:412718
  KEYWORDS   synthetic construct.
  SOURCE     synthetic construct.
  ORGANISM   artificial sequence.
  REFERENCE  1 (bases 1 to 791)
  AUTHORS    Ceska,M., Krizek,F., Hoegenauer,G. and Besemer,J.
  TITLE      Plasmids and bacteria strains, and method of preparing the same
  JOURNAL    Patent: EP 0084522-A 1 27-JUL-1983;
              SANDOZ AG; SANDOZ-PATENT-GMBH; SANDOZ-ERRINDUNGEN
              Vervalungseseelschaft m.b.H
  FEATURES   Location/Qualifiers
              1..791
              /organism="synthetic construct"
              /db_xref="taxon:32630"
BASE COUNT      206 a      179 c      193 g      174 t      39 others
ORIGIN

alignment_scores:
  Quality:      30.00      Length:      30
  Ratio:        1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x A06082 ..
```


Align seg 1/1 to: A06082 from: 1 to: 791

198 ThCysasnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
|||||
694 ACTGTATATGAGGACCGACCAATCTGACACATATATCTCAGGGAATA 743
214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
|||||
744 TCAATCAAAAGTTAAGAGCGACATATTTTCAGACTATCAG 783

seq_name: gb_pat:A06083

seq_documentation_block:

LOCUS A06083 791 bp DNA linear PAT 25-MAY-1993
DEFINITION PE8620 HindII-fragment (Reverse complement).
ACCESSION A06083
VERSION A06083.1 GI:411208
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 791)
AUTHORS Ceska,M., Kricek,F., Hoegenauer,G. and Besemer,J.
TITLE Plasmids and bacteria strains, and method of preparing the same
JOURNAL Patent: EP 0084522-A 2 27-JUL-1983;
SANDOZ AG; SANDOZ-PATENT-GMBH; SANDOZ-ERFINDUNGEN
Verwaltungsgerichtsschaft m.b.H

FEATURES
Source Location/Qualifiers

1..791 /organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 174 a 193 c 179 g 206 t 39 others
ORIGIN

alignment_scores:
Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A06083/rev ..

Align seg 1/1 to reverse of: A06083 from: 1 to: 791

198 ThCysasnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
|||||
98 ACTGTATATGAGGACCGACCAATCTGACACATATATCTCAGGGAATA 49
214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
|||||
48 TCAATCAAAAGTTAAGAGCGACATATTTTCAGACTATCAG 9

seq_name: gb_pat:A02701

seq_documentation_block:

LOCUS A02701 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production.
ACCESSION A02701
VERSION A02701.1 GI:344645
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 795)
AUTHORS
TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F
JOURNAL Patent: WO 8302456-A 1 21-JUL-1983;
FEATURES Location/Qualifiers

Source 1..795 /organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..795

BASE COUNT 207 a /note="enterotoxoid and antigens production"
ORIGIN 180 c 193 g 174 t 41 others

alignment_scores:
Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A02701 ..

Align seg 1/1 to: A02701 from: 1 to: 795

198 ThCysasnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
|||||
694 ACTGTATATGAGGACCGACCAATCTGACACATATATCTCAGGGAATA 743
214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
|||||
744 TCAATCAAAAGTTAAGAGCGACATATTTTCAGACTATCAG 783

seq_name: gb_pat:A02702

seq_documentation_block:

LOCUS A02702 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production (reverse complement).
ACCESSION A02702
VERSION A02702.1 GI:344646
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 795)
AUTHORS
TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F
JOURNAL Patent: WO 8302456-A 2 21-JUL-1983;
FEATURES Location/Qualifiers

Source 1..795 /organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature 1..795 /note="enterotoxoid and antigen production"
BASE COUNT 174 a 193 c 180 g 207 t 41 others
ORIGIN

alignment_scores:
Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A02702/rev ..

Align seg 1/1 to reverse of: A02702 from: 1 to: 795

198 ThCysasnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
|||||
102 ACTGTATATGAGGACCGACCAATCTGACACATATATCTCAGGGAATA 53
214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
|||||
52 TCAATCAAAAGTTAAGAGCGACATATTTTCAGACTATCAG 13

seq_name: gb_pat:AF452584

seq_documentation_block:

LOCUS AF452584 777 bp DNA linear BCT 14-JAN-2002
DEFINITION Vibrio cholerae strain 1322-69 CtxA (ctxA) gene, complete cds.
ACCESSION AF452584

VERSION AF452584.1 GI:18141148
KEYWORDS
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1.777
/organism="Vibrio cholerae"
/strain="1322-69"
/db_xref="taxon:666"
/note="identical sequence found in Vibrio cholerae strains
571-88 (serogroup O105) and 203-93 (serogroup O141)
serogroup: O37"
1.777
/gene="ctxA"
1.777
/gene="ctxA"
CDS 1.777
/transl_table=1
/product="CtxA"
/protein_id="AA160525.1"
/db_xref="GI:18141149"
/translation="MWKIIFFVFIFLSFSFYANDDKLYRADSRPPDEIKOSGGLMPRG
QNEYFDRTGTONINILYDHAFCGTGFRDDGVSTISLSRAHLVGOTILSGHSTYY
IYVITAPRMFVNVDVLGAYSPHDEQVSAIGTIPYQIYGMVRYHFGVLDLOLHRN
RGYRDRTYSLNDIAPADGGLAGFPRRAWREREPWHHAPPGCGNAPRSSMSWTCD
EKTQSLGVKFLDEYOSKVKRQIFSSGYQSDIDYHNRIKDEL"
BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN
alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AF452584 ..
Align seq 1/1 to: AF452584 from: 1 to: 777
28 AsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHi 44
|||||
136 AATGAGTACTTTGACGAGGACTCAATGATATCAACCTTTATGATCA 185
44 sAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
186 TGCAGAAGAGAACTCAGACGGGATTGTAGG 216
seq_name: gb_ba:AF463400
seq_documentation_block:
LOCUS AF463400 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 571-88 CtxA (ctxA) gene, complete cds.
ACCESSION AF463400
VERSION AF463400.1 GI:18448888
KEYWORDS
SOURCE Vibrio cholerae.

ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1.777
/organism="Vibrio cholerae"
/strain="571-88"
/db_xref="taxon:666"
/note="serogroup: O105"
1.777
/gene="ctxA"
1.777
/gene="ctxA"
CDS 1.777
/transl_table=1
/product="CtxA"
/protein_id="AA169944.1"
/db_xref="GI:18448889"
/translation="MWKIIFFVFIFLSFSFYANDDKLYRADSRPPDEIKOSGGLMPRG
QNEYFDRTGTONINILYDHAFCGTGFRDDGVSTISLSRAHLVGOTILSGHSTYY
IYVITAPRMFVNVDVLGAYSPHDEQVSAIGTIPYQIYGMVRYHFGVLDLOLHRN
RGYRDRTYSLNDIAPADGGLAGFPRRAWREREPWHHAPPGCGNAPRSSMSWTCD
EKTQSLGVKFLDEYOSKVKRQIFSSGYQSDIDYHNRIKDEL"
BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN
alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AF463400 ..
Align seq 1/1 to: AF463400 from: 1 to: 777
28 AsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHi 44
|||||
136 AATGAGTACTTTGACCGAGGACTCAATGATATCAACCTTTATGATCA 185
44 sAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
186 TGCAGAAGAGAACTCAGACGGGATTGTAGG 216
seq_name: gb_ba:AF463401
seq_documentation_block:
LOCUS AF463401 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 203-93 CtxA (ctxA) gene, complete cds.
ACCESSION AF463401
VERSION AF463401.1 GI:18448890
KEYWORDS
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.

TITLE Genetic Analysis of the Virulence Regions, CTX ϕ prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Li, M., Chen, Y., Kotelishvili, M., Stine, O. C., Morris, J. G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.

TITLE Direct Submission

JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA

FEATURES
source
1. 777
/organism="Vibrio cholerae"
/strain="203-93"
/db_xref="taxon:666"
/note="serogroup: O141"
1. 777
/gene="ctxA"
1. 777
/gene="ctxA"
/codon_start=1
/transl_table=1
/product="CtxA"
/protein_id="AA169945.1"
/db_xref="GI:18448891"
/translation="MKILVFVFILSSFSYANDKLYRADSRPPDEIKOSGGLMGRG
QNEYFDRTQMNIINLYDHARGTGTGVRHDDGVSTISLSRSHLVGOTILSGHSTYY
IYVIATAPMFMFVNDVLAGYSPHDEQVYSLGIPYSGIYRWVHFVLEQLHRN
RGYRDRIYENLDIAPADGGLAGFPPEHRAWREPEWIIHAPPGCGNAPRSSMSNTCD
EKTQSLGVKFLDEYQSKVKRQIFSGVQSDIDTNRKDEL"

BASE COUNT 243 a 129 c 169 g 236 t

ORIGIN

alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AF463401 ..

Align seg 1/1 to: AF463401 from: 1 to: 777

28 ASNCIUTYRPhEaSPARGlyThrGlnMeTasnlIeasnlEutyRAspH 44
|||||
136 AATGAGTACTTTGACCGAGGACTCAATGAAATCAACCTTTATGATCA 185
44 SAlaRgGlyThrGlnThrGlyPheValArg 54
|||||
186 TGCAGAGGAGACTCAGACGGGATTGTAGG 216

seq_name: gb_ba:VIBCTXABA

seq_documentation_block:
LOCUS VIBCTXABA 1372 bp DNA linear BCT 01-FEB-2000
DEFINITION Vibrio cholerae ctxA gene and ctxB gene for cholera toxins,
complete cds.
ACCESSION D30052
VERSION D30052.1 GI:487332
KEYWORDS ADP-ribosyltransferase; cholera toxin.
SOURCE Vibrio cholerae (sub-species O37, strain S7) (library: plasmid)
DNA, clone PK310.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 1372)
AUTHORS Yamamoto, K., Do, V. G., Xu, M., Iida, T., Mawatani, T., Albert, M. J. and
Honda, T.
TITLE Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio cholerae
strains 854 (O139-bengal) and S7 (O37) from two outbreaks
JOURNAL Unpublished (1994)

REFERENCE 2 (bases 1 to 1372)

AUTHORS Yamamoto, K.

JOURNAL Direct Submission

COMMENT Submitted (29-APR-1994) Kolchiro Yamamoto, Osaka University,
Institute for Microbial Diseases, Department of Bacterial
Infections, 3-1 Yamadaoka, Suita, Osaka 565, Japan
(Tel:06-879-4066)

Submitted (29-APR-1994) to DDBJ by:
Kolchiro Yamamoto
Institute for Microbial Diseases
Osaka University
3-1 Yamadaoka
Suita, Osaka 565
Japan
Phone: 06-879-4066
Fax: 06-879-8277.

FEATURES
source
1. 1372
/organism="Vibrio cholerae"
/strain="S7"
/sub_species="O37"
/db_xref="taxon:666"
/clone_lib="Plasmid"
113. 889
/gene="ctxA"
113. 889
/gene="ctxA"
/codon_start=1
/transl_table=1
/product="cholera toxin"
/protein_id="BA06288.1"
/db_xref="GI:808898"
/translation="MKILVFVFILSSFSYANDKLYRADSRPPDEIKOSGGLMGRG
QNEYFDRTQMNIINLYDHARGTGTGVRHDDGVSTISLSRSHLVGOTILSGHSTYY
IYVIATAPMFMFVNDVLAGYSPHDEQVYSLGIPYSGIYRWVHFVLEQLHRN
RGYRDRIYENLDIAPADGGLAGFPPEHRAWREPEWIIHAPPGCGNAPRSSMSNTCD
EKTQSLGVKFLDEYQSKVKRQIFSGVQSDIDTNRKDEL"

BASE COUNT 449 a 217 c 269 g 437 t

ORIGIN

sig_peptide
/gene="ctxA"
886. 1260
/gene="ctxB"
886. 1260
/gene="ctxB"
/codon_start=1
/transl_table=1
/product="cholera toxin"
/protein_id="BA06288.1"
/db_xref="GI:808900"
/translation="MIKKEGVFVILSSAVAHGTPTDLCAGYHNTQIHTLNDK
IISYRESLGNREMAITTFKNGATFPQVEVPGSCHIDSQKAIERMKDTLRIVLFEAK
VEKLCVNRNKTPTPAIAAISMAN"
886. 948
/gene="ctxB"

alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VIBCTXABA ..

Align seg 1/1 to: VIBCTXABA from: 1 to: 1372

28 ASNCIUTYRPhEaSPARGlyThrGlnMeTasnlIeasnlEutyRAspH 44
|||||
248 AATGAGTACTTTGACCGAGGACTCAATGAAATCAACCTTTATGATCA 297
44 SAlaRgGlyThrGlnThrGlyPheValArg 54
|||||

298 TGCAGAGAACTCAGACGGGATTGTTAGG 328

seq_name: gb_pat:A16422

seq documentation block:

LOCUS A16422 723 bp DNA linear PAT 06-JUL-1995

DEFINITION A subunit of cholera toxin CT.

ACCESSION A16422

VERSION A16422.1 GI:641010

KEYWORDS

SOURCE

ORGANISM

Vibrio cholerae.

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE

1 (bases 1 to 723)

IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN

LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES

Patent: WO 9313202-A 45 08-JUL-1993;

FEATURES

source

1..723

Location/Qualifiers

/organism="Vibrio cholerae"

/db_xref="taxon:666"

gene

1..723

/gene="CT subunit A"

CDS

<1..723

/gene="CT subunit A"

/codon_start=1

/transl_table=11

/protein_id="CAA01281.1"

/db_xref="GI:641011"

/translation="NDKTIKRDSPPEIKOSGLMPGSEFYEDKTKOMININLYD

ARGTGTGFVRHDDGYVSTISLSAHVGGTTLISGHSYIYIVATAPMFWNDVLG

AVSPHDEQEVSAIGIPYSQIYGMVRFVGLDGLHNRGVRDRYSNLDIAAAD

GYGLGFPPEHRAAREEPIWIIHAPPCGNAPRSSISNCDERTQSIGVFLDEYQSKV

KRIISGYSQSDIDTNRIRDEL"

BASE COUNT 227 a 125 c 163 g 208 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A16422 ..

Align seg 1/1 to: A16422 from: 1 to: 723

29 GluTyrPheAsparGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45

|||||

85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATGATCATGC 134

|||||

45 aArgGlyThrGlnThrGlyPheValArg 54

|||||

135 AAGAGAACTCAGACGGGATTGTTAGG 162

seq_name: gb_pat:AR118598

seq documentation block:

LOCUS AR118598 723 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6149919.

ACCESSION AR118598

VERSION AR118598.1 GI:14100508

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

1 (bases 1 to 723)

AUTHORS

Domenighini, M., Rappoli, R., Pizsa, M. and HOL, W.

TITLE

Immunogenic detoxified mutants of cholera toxin and of the toxin

LT, their preparation and their use for the preparation of vaccines

Patent: US 6149919-A 7 21-NOV-2000;

JOURNAL

FEATURES

source

1..723

Location/Qualifiers

/organism="unknown"

BASE COUNT 227 a 125 c 163 g 208 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AR118598 ..

Align seg 1/1 to: AR118598 from: 1 to: 723

seq_name: gb_ba:VIBCTXA

seq documentation block:

LOCUS VIBCTXA 777 bp DNA linear BCT 26-APR-1993

DEFINITION Vibrio cholerae enterotoxin A1 peptide (ctxA) gene, 5' end.

ACCESSION K02679

VERSION K02679.1 GI:155159

KEYWORDS

ADP-ribosyltransferase; enterotoxin.

V. cholerae (El Tor strain 62746) DNA.

SOURCE

Vibrio cholerae

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE

1 (bases 1 to 777)

Lockman, H.A., Galen, J.E. and Kaper, J.B.

Vibrio cholerae enterotoxin genes: Nucleotide sequence analysis of

DNA encoding ADP-ribosyltransferase

J. Bacteriol. 159, 1086-1089 (1984)

JOURNAL

85006737

COMMENT

This sequence is reported [1] to be identical to that for strains

El Tor 2125 and 569B. The 'ttttag' sequence repeated four times

beginning at position 38 and once more at 101 is correlated with

the amount of enterotoxin produced in various strains.

FEATURES

source

1..777

Location/Qualifiers

/organism="Vibrio cholerae"

/db_xref="taxon:666"

142..195

/note="enterotoxin signal peptide"

142..>777

/note="enterotoxin prepeptide"

/codon_start=1

/transl_table=11

/protein_id="AAA27514.1"

/db_xref="GI:155160"

/translation="MWKIIFFVFISSFSYVANDDKLYRADSRPPEIKOSGLMPRG

OSEYEDRSTOMKININLYDARGTGTGFVRHDDGYVSTISLSAHVGGTTLISGHSYI

IVVATAPMFWNDVLGAVSPHDEQEVSAIGIPYSQIYGMVRFVGLDGLHNRGVRDRYS

NDIAAADGYGLGFPPEHRAAREEPIWIIHAPPCGNAPRSS"

mat_peptide

196..777

/product="enterotoxin subunit A1"

BASE COUNT 237 a 119 c 164 g 257 t

ORIGIN

189 bp upstream of NdeI site.

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VIBCTXA

Align seg 1/1 to: VIBCTXA from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
280 GAGTACTTTGACGAGCTACTCAATGATATACACCTTTATGATCATGCG 329
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
330 AAGAGGAACTCAGACGCGATTGTAGG 357

seq_name: gb_pat:E00132

seq_documentation_block: 1148 bp DNA linear PAT 29-SEP-1997
LOCUS E00132
DEFINITION DNA coding of cholera toxin.
ACCESSION E00132
VERSION E00132.1 GI:2168431
KEYWORDS JP 1983222033-A/1.
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1148)
AUTHORS Arufuocoru, N. and Do, B.M.
TITLE SUBUNIT A AND B OF DNA ARRANGEMENT, RECOMBINED DNA, CHOLERA TOXIN
AND MEDICINE
PATENT: JP 1983222033-A 1 23-DEC-1983;

JOURNAL SMITH KLINE RIT

COMMENT OS cholera(Vibrio)
PN JP 1983222033-A/1
PD 23-DEC-1983
PF 23-MAY-1983 JP 1983091416
PR 24-MAY-1982 US 82 381083
PI ARUFUOCORU NIJIERU, DO BURUDO MISHIERU
PC A61K39/106,A61K39/00,C07H21/04,C12N15/00;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=ATCC 39050;
FH Key Location/Qualifiers

FT CDS 1..777
FT /product='subunit A of cholera toxin' FT
FT mat_peptide 1..774
FT /product='subunit A of cholera toxin' FT CDS
FT 774..1148
FT /product='subunit B of cholera toxin' FT
FT mat_peptide 774..1145
FT /product='subunit B of cholera toxin' FT
FEATURES
source Location/Qualifiers
1..1148
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 382 a 184 c 235 g 347 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x E00132

Align seg 1/1 to: E00132 from: 1 to: 1148

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGCTACTCAATGATATACACCTTTATGATCATGCG 188

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGAGGAACTCAGACGCGATTGTAGG 216

seq_name: gb_ba:VCTOXA21

seq_documentation_block:

LOCUS VCTOXA21 1362 bp DNA linear BCT 23-SEP-1991
DEFINITION V.cholerae genes for toxin proteins A and B, strain 2125.
ACCESSION X58786
VERSION X58786.1 GI:48420
KEYWORDS cholera toxin.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae.

REFERENCE 1 (bases 1 to 1362)
AUTHORS Dams,E.
TITLE Direct Submission

JOURNAL

Belgium
2 (bases 1 to 1362)
Dams,E., De Wolf,M. and Dierick,W.
Correction of the cholera toxin nucleotide sequence of the Vibrio
cholerae eltor strain 2125
Unpublished

FEATURES
source Location/Qualifiers
1..1362
/organism='Vibrio cholerae'
/strain='eltor 2125'
/db_xref='taxon:666'
/clone='PRT10824'
72..92
/rpt_type='TANDEM'
110..121
136..142
169..945
/codon_start=1
/transl_table=11
/product='cholera toxin A protein (CTA)'
/protein_id='CAA41592.1'
/db_xref='GI:48421'
/db_xref='SWISS-PROT:P01555'
/translation='MKITFVFPIFLSSFSYANDKLYRADSRPPPEIKQSGLMRG
QSEYEDRGTMNINLYDHARGTQGVVRHDDGVSTLSLSHLVGTLSGHSTY
IYVATAPMFMFVNDVLGAYSPHPDQEVSLGIGYSGIYGRVHFGVLEQLHRN
RGVDRRYENLDIAPADGYGAGPPEHRAMREEPFMIHHAPGCGNADRRSSMNTCD
EKQSGVKEFLEDEYQSKVKRQIFSGVQSDIDFNRKIDEL'
169..222
223..942
/product='cholera toxin A protein (CTA)'
942..1004
942..1316
/codon_start=1
/transl_table=11
/product='cholera toxin B protein (CTB)'
/protein_id='CAA41593.1'
/db_xref='GI:48422'
/db_xref='SWISS-PROT:P01556'
/translation='MIKLFQVGFVTLSSAYAHGTPONTDLCAEYHNTQIYTLNDK
IFSYPESLAGKREMAITPFGKCAIFQVEVPSCGHIDSOKKAIERMKDTRIIVLTFAK
VERLCVNNKTRPHAIATAISMAN'
1005..1313
/product='cholera toxin B protein (CTB)'

BASE COUNT 455 a 205 c 276 g 426 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VCTOXB21 ..

Align seg 1/1 to: VCTOXB21 from: 1 to: 1362

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 |||
 307 GAGTACTTTCACCGAGTACTCAATGATATCAACCTTATGATCATGCG 356
 45 aArgGlyThrGlnThrGlyPheValArg 54
 |||
 357 AAGAGAACCTCAGACGGGATTGTAGG 384

seq_name: gb_ba:VIBCTYXAB

seq_documentation_block:

LOCUS VIBCTYXAB 1369 bp DNA linear BCT 01-FEB-2000
 DEFINITION Vibrio cholerae ctxA gene and ctxB gene for cholera toxins,
 complete cds.

ACCESSION D30053

VERSION D30053.1 GI:487333

KEYWORDS ADP-ribosyltransferase; cholera toxin.

SOURCE Vibrio cholerae (sub-species O139-Bengal, strain 1854) (library:
 plasmid) DNA, clone pKY340.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 1369)
 Yamamoto, K., Do, Y.G., Xu, M., Iida, T., Miwatani, T., Albert, M.J. and
 Honda, T.

TITLE

Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio cholerae
 strains 854 (O139-Bengal) and S7 (O37) from two outbreaks

JOURNAL Unpublished (1994)

REFERENCE 2 (bases 1 to 1369)

AUTHORS

Yamamoto, K.

TITLE

Direct Submission
 Submitted (29-APR-1994) Koichiro Yamamoto, Osaka University,
 Institute for Microbial Diseases, Department of Bacterial
 Infections; 3-1 Yamadaoka, Suita, Osaka 565, Japan
 (Tel:06-879-4066)

COMMENT Submitted (29-APR-1994) to DDBJ by:

Koichiro Yamamoto
 Institute for Microbial Diseases
 Osaka University
 3-1 Yamadaoka
 Suita, Osaka 565

Japan
 Phone: 06-879-4066
 Fax: 06-879-8277.

FEATURES Location/Qualifiers

source 1..1369
 /organism="Vibrio cholerae"
 /strain="1854"
 /sub_species="O139-Bengal"
 /db_xref="taxon:666"
 /clone_id="plasmid"
 110..886
 /gene="ctxA"
 110..886
 /gene="ctxA"
 /codon_start=1
 /transl_table=11
 /product="cholera toxin"
 /protein_id="BAA06290.1"
 /db_xref="GI:808901"
 /translation="WKIIFVFIFLSEFSYANDKLYRADSRPPDEIKQSGGLMPRG
 OSEYFRGTQMINIVLYDHARGTQTGFVRHDDQVSTSLRSALHVGQTILSGHSTY
 IYVATAPNFMVNDVLYGASPHDEQESALGIPYSQIGMYRVHVGVLDEQLHRN
 RGYRDYYSNLDIAPADGYGLAGFPPEHRAREPPIHHAPGCGNAPRRSSMNTCD
 EKTOSLGYKFLDEYQSKVRQIFSGYQSDIDIHNRKDEL"
 110..163
 /gene="ctxA"
 883..1257

CDS

/gene="ctxB"
 883..1257
 /gene="ctxB"
 /codon_start=1
 /transl_table=11
 /product="cholera toxin"
 /protein_id="BAA06291.1"
 /db_xref="GI:808902"
 /translation="MIKIKFGVFVLLSAYAHGTPTNIDLCAEYHNTOIYTLNDR
 LFSYRESIAGKREMAITFEKCAIFQVEVPQSQIHDSQKAIERKKDPLRIAYLLEAK
 VEKLCVNNKTPHAIATISMAN"
 883..945
 /gene="ctxB"

sig_peptide

BASE COUNT 449 a 214 c 269 g 437 t
 ORIGIN

alignment_scores:

Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTYXAB ..

Align seg 1/1 to: VIBCTYXAB from: 1 to: 1369

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 |||
 248 GAGTACTTTCACCGAGTACTCAATGATATCAACCTTATGATCATGCG 297
 45 aArgGlyThrGlnThrGlyPheValArg 54
 |||
 298 AAGAGAACCTCAGACGGGATTGTAGG 325

seq_name: gb_ba:VCTOXB56

seq_documentation_block:

LOCUS VCTOXB56 1397 bp DNA linear BCT 30-JUN-1993
 DEFINITION V. cholerae genes for toxin proteins A and B, strain 569B.

ACCESSION X58785

VERSION X58785.1 GI:48888

KEYWORDS cholera toxin.

SOURCE Vibrio cholerae.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 1 (bases 1 to 1397)

REFERENCE

Dams, E.

AUTHORS

Direct Submission

TITLE

Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
 RUCA University Antwerp, Gruenenburgerlaan 171, B-2020 Antwerp,
 Belgium

JOURNAL

Biochim. Biophys. Acta 1090 (1), 139-141 (1991)

MEDLINE 91355224

FEATURES

Location/Qualifiers
 1..1397
 /organism="Vibrio cholerae"
 /strain="classical strain 569B"
 /db_xref="taxon:666"
 /clone="PJM17"
 72..127
 /rpt_type=TANDEM
 /rpt_type=TANDEM
 145..156
 171..177
 204..980
 /codon_start=1
 /transl_table=11
 /product="cholera toxin A protein (CTA)"
 /protein_id="CAA41590.1"

sig_peptide

gene
 883..1257


```
/db_xref="GI:48889"
/db_xref="SWISS-PROT:P01555"
/translation="WVKIIFVFIFLSSFSYANDKLYRADSRPPEIKQSGGLMPRG
QSEYFDRTQMINILYDHARGTQTFVRHDDGVSTSLRSANHLVGOTILSGHSTYY
IYVATAPNMFNVNDVLGAYSPHDEQVSAAGLIPYQIYGMRYVHFGVLEQLHRN
RGYRDRYVSNLDIAPAADGYGAGPPEHRAAREPWIHHAPGCCGNAPRRSSMTNCD
EKTQSLGVKFLDEYQSKVKRQIFSGYQSDIDITHNKIDEL"
204..257
sig_peptide
258..977
mat_peptide
sig_peptide
977..1039
/product="cholera toxin A protein (CTA)"
977..1351
/codon_start=1
/transl_table=11
/product="cholera toxin B protein (CTB)"
/db_xref="GI:48890"
/db_xref="GI:48890"
/db_xref="SPTREMBL:Q57193"
/translation="MIKLKFGVFTVLLSSAYAHGTPONITDCAEYHNTQIHTLNDK
IESYVESLAGKREMAITPEKNGATFOVEPSSQIHDSQKKAERKKDILRIAYLLEAK
VERLCVWNKTRPHAIATISMAN"
1040..1348
mat_peptide
1040..1348
/product="cholera toxin B protein (CTB)"
BASE COUNT 460 a 207 c 281 g 449 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VCTOXAB56 ..
Align seg 1/1 to: VCTOXAB56 from: 1 to: 1397

29 G|U|T|Y|R|H|E|A|S|P|A|G|G|Y|T|H|G|M|E|C|A|S|N|L|E|A|S|N|E|U|T|Y|R|A|S|P|H|S|A|L 45
|||||
342 GAGTACTTTCACGAGCTACTCAATGATATCAACCTTATGATCATGC 391
45 AARGGLYTHrGlnThrGlyPheValArg 54
|||||
392 AAGAGAACTCAGACGCGATTGTGTAGG 419

seq_name: gb_ba:VCCTX

seq_documentation_block:
LOCUS VCCTX 2020 bp DNA linear BCT 30-MAR-1995
DEFINITION Vibrio cholera toxin (ctx) operon DNA sequence from strain 2125.
ACCESSION X00171
VERSION X00171.1 GI:48347
KEYWORDS overlapping genes: signal peptide; tandem repeat.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 2020)
AUTHORS Mekalanos,J.U., Swartz,D.J., Pearson,G.D., Harford,N., Groyne,F.
and de Wilde,M.
TITLE Cholera toxin genes: nucleotide sequence, deletion analysis and
vaccine development
JOURNAL Nature 306 (5943), 551-557 (1983)
MEDLINE 84068199
FEATURES
Location/Qualifiers
source
1..2020
/organism="Vibrio cholerae"
/db_xref="taxon:666"
misc_feature
419..425
/note="tandem repetitive sequence"
misc_feature
426..432
/note="tandem repetitive sequence"
misc_feature
433..439
/note="tandem repetitive sequence"
```

```
promoter
457..468
/note="35 region"
promoter
483..489
/note="10 region"
RBS
504..508
/note="Shine-Dalgarno sequence"
RBS
507..510
/note="Shine-Dalgarno sequence"
CDS
516..1292
/note="ctx A"
/codon_start=1
/transl_table=11
/product="ctx A"
/db_xref="GI:48348"
/db_xref="SWISS-PROT:P01555"
/translation="WVKIIFVFIFLSSFSYANDKLYRADSRPPEIKQSGGLMPRG
QSEYFDRTQMINILYDHARGTQTFVRHDDGVSTSLRSANHLVGOTILSGHSTYY
IYVATAPNMFNVNDVLGAYSPHDEQVSAAGLIPYQIYGMRYVHFGVLEQLHRN
RGYRDRYVSNLDIAPAADGYGAGPPEHRAAREPWIHHAPGCCGNAPRRSSISMTNCD
EKTQSLGVKFLDEYQSKVKRQIFSGYQSDIDITHNKIDEL"
516..524
sig_peptide
522..1151
mat_peptide
1152..1289
/product="ctx A1 polypeptide"
gene
1289..1663
/note="ctx B"
/gene="ctx B"
1289..1351
sig_peptide
1289..1351
mat_peptide
1289..1663
/note="ctx B"
/gene="ctx B"
1289..1663
CDS
1289..1663
/gene="ctx B"
/product="ctx B"
1352..1660
mat_peptide
1352..1660
/product="ctx B"
BASE COUNT 602 a 352 c 446 g 620 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VCCTX ..
Align seg 1/1 to: VCCTX from: 1 to: 2020

29 G|U|T|Y|R|H|E|A|S|P|A|G|G|Y|T|H|G|M|E|C|A|S|N|L|E|A|S|N|E|U|T|Y|R|A|S|P|H|S|A|L 45
|||||
654 GAGTACTTTCACGAGCTACTCAATGATATCAACCTTATGATCATGC 703
45 AARGGLYTHrGlnThrGlyPheValArg 54
|||||
704 AAGAGAACTCAGACGCGATTGTGTAGG 731

seq_name: gb_pat:E03512

seq_documentation_block:
LOCUS E03512 2020 bp DNA linear PAT 29-SEP-1997
DEFINITION Cholera toxin gene.
ACCESSION E03512
VERSION E03512.1 GI:2171728
KEYWORDS JP 199209488-A/17.
SOURCE synthetic construct.
```



```

ORGANISM      synthetic construct
REFERENCE     1 (bases 1 to 2020)
AUTHORS      Shirai,H.
TITLE        PRIMER FOR AMPLIFICATION OF GENE
JOURNAL      Patent: JP 1992099488-A 17 31-MAR-1992;
              UNITIKA LTD
              PN JP 1992099488-A/17
              PD 31-MAR-1992
              PP 20-AUG-1990 JP 1990219523
              PI SHIRAI HIROMASA
              PC C12N15/11,C12Q1/68;
              CC strandedness: Double;
              CC topology: Linear;
              FH Key
              FT misc-feature 1..2020
               /note="cholera toxin gene".
FEATURES
  source      1..2020
               /organism="synthetic construct"
               /db_xref="taxon:32630"
BASE COUNT   602 a 352 c 446 g 620 t
ORIGIN
alignment_scores:
  Quality:    26.00      Length:    26
  Ratio:      1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-528-682-1 x E03512 ..
Align seg 1/1 to: E03512 from: 1 to: 2020
29 GtutyrpheasparglythrGlnMetasnleasnleutyrrasphisal 45
|||||
654 GAGTACTTTGACCGAGTACTCAATGAATGATACCTTTATGATCATGC 703
45 aargglythrGlnthrGlyphevalarg 54
|||||
704 AAGAGGAACTCAGACGGAGATTGTAGG 731
seq_name: gb_ba:AF175708
seq_documentation_block:
LOCUS      AF175708      3480 bp      DNA      linear      BCT 15-NOV-1999
DEFINITION Vibrio cholerae accessory cholera enterotoxin (ace), zonular
occludens toxin (zot), cholera enterotoxin A-subunit (ctxA), and
cholera enterotoxin B-subunit (ctxB) genes, complete cds.
ACCESSION  AF175708
VERSION     AF175708.1 GI:5764645
KEYWORDS
SOURCE      Vibrio cholerae.
ORGANISM    Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE   1 (bases 1 to 3480)
AUTHORS    Shin,H.J., Park,Y.C. and Kim,Y.C.
TITLE      Cloning and nucleotide sequence analysis of the virulence gene
cassette from Vibrio cholerae KNIH002 isolated in Korea
JOURNAL    Misaiumuhag Ho1j1 35 (3), 205-210 (1999)
REFERENCE   2 (bases 1 to 3480)
AUTHORS    Shin,H.J., Park,Y.C. and Kim,Y.C.
TITLE      Direct Submission
JOURNAL    Submitted (04-AUG-1999) School of Life Sciences, Chungbuk National
University, Kaesin-Dong san 48, Cheongju, Chungbuk-do 361-763,
South Korea
FEATURES
  source      1..3480
               /location/Qualifiers
               /organism="Vibrio cholerae"
               /strain="KNIH002"
               /db_xref="taxon:666"

```

```

gene        /country="South Korea"
            287..577
            /gene="ace"
CDS         287..577
            /gene="ace"
            /codon_start=1
            /product="accessory cholera enterotoxin"
            /protein_id="AAD51357.1"
            /db_xref="GI:5764646"
            /translation="MAMDLYDLIDGFTWLVYKGLIMIESKIFVIOFWEWSQKV
            IDMFITYPLIDQALDLPQYSGFLFGLDQALVIALQALMTFRALNALM."
            574..1773
            /gene="zot"
            574..1773
            /gene="zot"
            /note="cholera enterotoxin"
            /codon_start=1
            /transl_table=1
            /product="zonular occludens toxin"
            /protein_id="AAD51358.1"
            /db_xref="GI:5764647"
            /translation="MSIFIHGAPSGSYKTSQALMRLPAIKSGRHITTYNGINLER
            MAKYLMKDVSDISIEFTDHPDGRLTMAKREWHARKDAFLFIDCGIWPRLVTN
            LKALDTPDVLVAEDRESFEVAFDMHRHGWIDICLTENIAKVNMMIREAEIGRHF
            NRAIVGLGAKFTLTTHDAANSQOMDSHALTRQVKRIPSTIFKMYASTITGKARDTMAG
            TALMKDRKILFLFGWVFLMFSSYFGLHDNPFTTGGNDATIESEQSEQSAATVGNV
            GSKAVAPASFGFCIGRICVODGEFVVGDERRLVNDIPYRGILWATGHIIYKDTLV
            PRTESGSPPTFLPASSTYKRVLPIDRNHPVPTPTFAQALMVEYKRGKGLPIKTEWDK
            KCLNSTF"
            1872..2648
            /gene="ctxA"
            1872..2648
            /gene="ctxA"
            /codon_start=1
            /transl_table=1
            /product="cholera enterotoxin A-subunit"
            /protein_id="AAD51359.1"
            /db_xref="GI:5764648"
            /translation="MWKIIFFVFIFLSSFSFYANDDKLYRADSPPEDEIKOSGLMPRG
            OSEVFRGROMNINLYDHARGOTGFVRRDQGVSTSISSRAHLYGOTILHSGHSTY
            IYVITAPRMENVNDVLCATSPHPDEQESYALGCIPTYSQIGMYRVHVGVDQJHRN
            RGYRDYYSNLDIAPADQYGLAGFPPEHRARKEPWHHAPPGCGNAPRSMSTCD
            EKTQSLGKVFLEIDQYQSKVKRQIFSGYQSDIDTHNRKDEL"
            2645..3019
            /gene="ctxB"
            2645..3019
            /gene="ctxB"
            /codon_start=1
            /transl_table=1
            /product="cholera enterotoxin B-subunit"
            /protein_id="AAD51360.1"
            /db_xref="GI:5764649"
            /translation="MIKIKGVFTVYLSSAAGHTPONITDCAEYHNTQIHTLNDK
            IFSYDSLAKREMAIITFKNGATFOVEPGSQHDSQKALDRKDKTLRIAYLLEAK
            VERLCVWNKRTPHALAAISMAM"
BASE COUNT   948 a 721 c 808 g 1003 t
ORIGIN
alignment_scores:
  Quality:    26.00      Length:    26
  Ratio:      1.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-528-682-1 x AF175708 ..
Align seg 1/1 to: AF175708 from: 1 to: 3480
29 GtutyrpheasparglythrGlnMetasnleasnleutyrrasphisal 45
|||||
2010 GAGTACTTTGACCGAGAACCTCAATGATATCAACCTTATGATCATGC 2059

```


45 aargglythrGlnThrGlyPheValArg 54
|||||
2060 AAGAGGACTCAGACGGGATTGTAGG 2087

seq_name: gb_ph:AF414369

seq_documentation_block:

LOCUS AF414369 4275 bp DNA linear PHG 02-OCT-2001

DEFINITION Vibrio phage CTX Ace (ace), Zot (zot), CtxA (ctxA), and CtxB (ctxB) genes, complete cds.

ACCESSION AF414369

VERSION AF414369.1 GI:15825396

KEYWORDS

SOURCE

ORGANISM

Vibrio phage CTX.

Viruses; ssDNA viruses; Inoviridae; Inovirus.

1 (bases 1 to 4275)

Bhattacharyya,T., Nandy,R.K. and Nair,G.B.

The entire core region of the ctx-phl (ctx-prophage) in VCE 232, an environmental strain of V. cholerae

Unpublished

2 (bases 1 to 4275)

Bhattacharyya,T., Nandy,R.K. and Nair,G.B.

Direct Submission

Submitted (27-AUG-2001) Microbiology, NICED, Calcutta, India

FEATURES

source

1. 4275

/organism="Vibrio phage CTX"

/specific_host="Vibrio cholerae strain VCE 232"

/db_xref="taxon:141904"

/note="prophage: core region"

1543..1833

/gene="ace"

1543..1833

/gene="ace"

/codon_start=1

/transl_table=1

/product="ace"

/protein_id="AAL09683.1"

/db_xref="GI:15825396"

/translation="MLMDTLVDMLIDGFTWLVIKGLIMIESKIFVIOFWMSQKV

IDMFTIYPLIQADIMLPPQYSGFLFGLDQALAIYQALMTRELRALNLI"

1830..3029

/gene="zot"

1830..3029

/gene="zot"

/codon_start=1

/transl_table=1

/product="zot"

/protein_id="AAL09684.1"

/db_xref="GI:15825400"

/translation="MSIFIHGAPGSYKTSGLMLRLPAKSGRIITNVNGLIER

MAKYLKMDVSDISIEFDTHDPGRSLTMAKRWMAKDAFLFIDECGIMPRLATN

LKALDTPDLVAEDRSEFEVAFDMHRRHGMDICLTPNIKVMHMEAEIGRHF

NRAVVGAGKFTLTHDANSNGSDSHALFTQVKRIPSPIKMYASTTGARDMAG

TALKDKRIILFEGMFLMFSYFGLDNFIIFGNGNATIESSESEQSATAGNAV

GSKAVAPASFEFCIGRLCVQDGFVTGDERILVDNDIPIYRGALWATGHILKDLTY

FEETESGVSPELFASSYRYKVLPLDPENHVPDTEAQAALWEVRRGLPIKTENDK

GLNSIF"

3128..3904

/gene="ctxA"

3128..3904

/gene="ctxA"

/codon_start=1

/transl_table=1

/product="ctxA"

/protein_id="AAL09681.1"

/db_xref="GI:15825397"

/translation="WKRIIFYFIFLSSFYANDKLYRADSRPPDEIKSGGLMPRG

QSEYFDKOTONILYDHARGTGTGVNHDGCVSTISLSAHLVGGTILSGSTYY

IYVATAPNMENVDVLAAYSHPDDEQVSAIGIPIYQIYGVFVHGVLDQQLHRN

RGYRDYYSNDIAPADGVLGAGFPREHRAWRBEPWIHAPPCGNGAPRSSMSNTCD

gene
EKTOSLGVFLDEYOSKVKROIFSGYQSDIDTHNRKIDEL"

3901..4275

/gene="ctxB"

3901..4275

/gene="ctxB"

/codon_start=1

/transl_table=1

/product="ctxB"

/protein_id="AAL09682.1"

/db_xref="GI:15825398"

/translation="WIKLFGVFTVLLSAYANHTPHNITATCAEYNTQIHTLNDK

IFSYTESLAGKREMAIIFTEKNGAFQVVPQSOHIDSKKAIERRKIDRLRAYLLEAK

VEKLCVMNNKTPHAIATASMAN"

BASE COUNT 1132 a 904 c 997 g 1242 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AF414369 ..

Align seg 1/1 to: AF414369 from: 1 to: 4275

29 GltUTrPheAspArgGlyThrGlnMetAsnIleAsnLeuTrpAspHisAl 45

|||||

3266 GAGTACTTGTACCGAGGACTCAATGAATCAACCTTATGATCATGTC 3315

45 aargglythrGlnThrGlyPheValArg 54

|||||

3316 AAGAGGACTCAGACGGGATTGTAGG 3343

seq_name: gb_ba:AE004224

seq_documentation_block:

LOCUS AE004224 10646 bp DNA linear BCT 31-JUL-2000

DEFINITION Vibrio cholerae chromosome I, section 132 of 251 of the complete

chromosome.

ACCESSION AE004224 AE003852

VERSION AE004224.1 GI:9655952

KEYWORDS

SOURCE

ORGANISM

Vibrio cholerae.

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

1 (bases 1 to 10646)

Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwin,M.L.,

Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D.,

Ernstlaeva,M.D., Vamathevan,J., Bass,S., Qin,H., Dragol,I.,

Sellers,P., McDonald,L., Uitterback,T., Fleischmann,R.D.,

Nierman,W.C. and White,O.

DNA sequence of both chromosomes of the cholera pathogen Vibrio

cholerae

Nature 406 (6795), 477-483 (2000)

2 (bases 1 to 10646)

Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwin,M.L.,

Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.A.,

Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D.,

Ernstlaeva,M.D., Vamathevan,J., Bass,S., Qin,H., Dragol,I.,

Sellers,P., McDonald,L., Uitterback,T., Fleischmann,R.D.,

Nierman,W.C., White,O., Salzberg,S.L., Smith,H.O., Colwell,R.R.,

Mekalanos,J.J., Venter,J.C. and Fraser,C.M.

Direct Submission

Submitted (14-JUN-2000) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

Location/Qualifiers

1..10646

/organism="Vibrio cholerae"

/strain="NJ9641"

gene /serotype="01"
/db_xref="taxon:666"
/chromosome="1"
/note="biotype: E1_Tor"
complement(215. .439)
/gene="VC1452"
complement(215. .439)
/note="similar to GP:2564354; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstC protein"
/protein_id="AAF94609.1"
/db_xref="GI:9655953"
/translation="MSLKPYTMDVYDSIEDLNNAALYRSQATDEIAHQYNLICD
KIIDQGIYNERLSPSLNPKALTEPSL"
complement(333. .907)
/gene="VC1453"
complement(533. .907)
/note="similar to GP:2564353; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstB1 protein"
/protein_id="AAF94610.1"
/db_xref="GI:9655954"
/translation="MMKLMVNMKSRFPVFGASHSEGVSKTAPYLIPVYGRKIRO
MKNDGQCLTEGLQHOEVKFTVSSDAMTRKLEOTAPVLTFFNEPDEPDSNVLIDY
QVCSLFDPNVPQKPLDKPKNT"
complement(885. .1964)
/gene="VC1454"
complement(885. .1964)
/note="similar to GP:2564352; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="RstA1 protein"
/protein_id="AAF94611.1"
/db_xref="GI:9655955"
/translation="MKQLFTLDELQDTNAPVFDYLANSPVASFRAHKSDLS
SLIMAPLPKPYDMARTEPEKLEILEYKOKNVAMMERLEFCLHVLGRSPMRDK
GLYGYENSHIDMSKYSNKHGVALGNNNTCYFOIEGVGRTVLEHTSIFELHMLD
LIGCSRLSIDLAVDDPFGHGFEGREYAKKAYSDAPRTAAGRAPNGGEVLEPPEKI
INSEFEVGSRESRIYRTYKNAOGLIDMHMRNREVELKMDP IDVLNTEGVFAGLCA
YASIIINSLPKVVTYKKROYALDIHSRIKMAROYKTLFDISKHFGDLERVFQALI
SKFIHDSLNPDPSTMKLIDETMGD"
2090. .2428
/gene="VC1455"
2090. .2428
/note="similar to GP:2564351; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="transcriptional repressor RstR"
/protein_id="AAF94612.1"
/db_xref="GI:9655956"
/translation="MKIKERLANORKAIKNTQAOAMADEIGISLTYSKYSEGGLPMT
ENLVKADALEISIDELCGRMATDENOEMLRLKIQOQDEDEOKAISMVLESLMRH
STKSLINHGA"
complement(3030. .3404)
/gene="VC1456"
complement(3030. .3404)
/note="similar to GB:K01170 SP:P01556 GB:D30052 GB:D30053
GB:X58785; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin, B subunit"
/protein_id="AAF94613.1"

gene /db_xref="GI:9655957"
/translation="MIKIKFGVFFTVLLSSAYHGTPONITDLCAEYHNTQIYTLNDK
IFSYTESLAGKREMAITTEKNGALFOVEVPGSIDSQKKAIEHMKDTRLRIAYLLEAK
VEKICVWNKTRPHAIAMISMAN"
complement(3401. .4177)
/gene="VC1457"
complement(3401. .4177)
/note="similar to SP:P01555 GB:D30052 GB:D30053 GB:K02679
GB:X58785; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin, A subunit"
/protein_id="AAF94614.1"
/db_xref="GI:9655958"
/translation="MWKLIIFVFETFLSEFSYANDDKLYRADSRPEDEIKOSGLMPRG
OSEYFDRGCTOKNINLYDARGTQGFVARDHGGVSTISLSAHLVQOTILSGHSTY
IYVATAPMFRNVDVIGAYSPPHDEQVSNALGGITPISQIYGMIRYHFGVLDEDLHRN
RGYRDYRYSNLIDIAPAADGYGLAGPPEHRAWRREPMIHAHPGCGNAPRSSMNTCD
EKTQSLGVKFLDEYQSKYKQIFSGYOSDIDITHNRIRDEL"
complement(4276. .5475)
/gene="VC1458"
complement(4276. .5475)
/note="similar to GB:M83563 SP:P38442 PID:155315;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="zona occludens toxin"
/protein_id="AAF94615.1"
/db_xref="GI:9655959"
/translation="MSIFIHGAGQSYKTSGLMWLRLLPAIKSGRHITTYNGINLER
MAKYLMKDVSIDISIEFIDTDHPDRLTWARPMHARKDAFLFIDCGRIWPRLTYN
LKALDTPDLVAEDRESEFEVAFDMRHMGDICTLPINAKVHMTREAEIYRHF
NRATYVGAKRFTITTHDANSNGDSHALTFQVKKIPSPITKMYASTPGKADPTMG
TALKDKRIITLFCGVPLMFSYSFGIHDNIFPLGGDARTIESQOSEQSKATGMNV
GSKAVAPASFGCIGRLCVQDGFVTVGERIRLVDNDIDPRGLMAGHNIYKOTLY
FEETESGVPTELFASSYRYKVLPLDPNHFVDPDTAAQALWVEVRKGLPIKTENDK
GLNSIF"
complement(5472. .5765)
/gene="VC1459"
complement(5472. .5765)
/note="similar to SP:P38441 GB:Z22569 PID:298062;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="accessory cholera enterotoxin"
/protein_id="AAF94616.1"
/db_xref="GI:9655960"
/translation="MMLMDTLYDMLIDGFTWLYIKGIMNIESKIFVIOFWMESOK
VIDMFTIYPLIQOALIDMLPOYSGFLFFGLDQALAIYQALMTFALRALNL"
complement(5762. .7045)
/gene="VC1460"
complement(5762. .7045)
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94617.1"
/db_xref="GI:9655961"
/translation="MAGSILMSLGHHRVLCVGLIFGLREGESESPFMRYFLTLTLF
LSPSYTASAINCDPMTWTSHQDLFGSPVIOVSULFDCMDIKDQDGFYWGSLSN
NGDYCKGLKRPRESQGVSPNMPDCDLSGASAEKCIYVCPGREGCVPLPSPSPSPV
DGLSSRSKSAFNQYTKRNOSEMASTLNHVSQVSHSQDMVOLNTFFHADRYESTAYN
NRLGQMBYLEIRIDVDTQREVRKADELYSRVAASVYLVSELVLAIDELKD
SLGTVVPPNPDPQNPPTPDDSSPWTGALNTISKILNTLETISQOQDTMTALSGRC
SNPERCQPIREAELETETAOONKONIMETKTIQSALQFSGSAVVSFCSYVEAFGY
NLCFDFSLSEHIAHTIRMIYLAAMAYIIAAILLFR"
complement(7056. .7304)
/gene="VC1461"
complement(7056. .7304)


```

/gene="VC1461"
/Note="similar to ; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="colonization factor"
/protein_id="AAF94618.1"
/db_xref="GI:9655962"
/translation="MESSKKKLNFKSTLSGVFLTSAPFANQALAADGLVTEVT
KTLGTSKDTVALGPLMGVGAIVLIVTGLIRKAK"
complement(7440..7823)
/gene="VC1462"
complement(7440..7823)
/Note="similar to GP:2564358; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="RatB2 protein"

```

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AE004224/rev ..

Align seg 1/1 to reverse of: AE004224 from: 1 to: 10646

```

29 GUTYrPheasparGlyThrGlnMetAsnIleAsnLeuTyraSPHISAI 45
|||||
4039 GAGTACTTTCACGAGTACTCAAAATATCACTTTATGATCATGCC 3990
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
3989 AAGAGCAATCAAGACGGATTGTTAGG 3962

```

seq_name: gb_ba:ECOELTA2

seq_documentation_block:

```

LOCUS      ECOELTA2              165 bp    DNA    linear    BCT 26-APR-1993
DEFINITION E.coli heat labile enterotoxin subunit a gene; a1/a2 junction.
ACCESSION  K00433
VERSION    K00433.1 GI:145829
KEYWORDS   enterotoxin.
SOURCE      Escherichia coli DNA.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE  1 (bases 1 to 165)
AUTHORS   Spierer,E.K., Kavanaugh,W.M., Dallas,W.S., Falkow,S.,
            Konigsberg,W.H. and Schaefer,D.E.
TITLE      Sequence homologues between a subunits of Escherichia coli and
            Vibrio cholerae enterotoxins
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 78, 50-54 (1981)
MEDLINE    81233767
COMMENT    see other loci beginning <ecoelt>.
FEATURES   location/Qualifiers
            source          1..165
                        /organism="Escherichia coli"
                        /db_xref="taxon:562"

```

BASE COUNT 59 a 30 c 35 g 41 t
 ORIGIN about 300 bp after seg 1.

alignment_scores:
 Quality: 24.00 Length: 24
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOELTA2 ..

Align seg 1/1 to: ECOELTA2 from: 1 to: 165

```

195 ThrGlyasprThrcysaanglGluThrGlnAsnLeuSerThrIleTyrl 211
|||||
55 ACAGGTACTACTTGTAAATGAGAGAGCCAGAAATCTGAGACAAATATATCT 104
211 uArgGluTyrcInSerIysVal 218
|||||
105 CAGGGAATATCAATCAAAAGTT 126

```

seq_name: gb_ba:ECOELT

seq_documentation_block:

```

LOCUS      ECOELT              774 bp    DNA    linear    BCT 25-JUN-1996
DEFINITION E.coli heat labile enterotoxin a 3' end and enterotoxin b (tox)
            gene, complete cds.
ACCESSION  J01646
VERSION    J01646.1 GI:145828
KEYWORDS   enterotoxin.
SOURCE      Escherichia coli (strain H10407) DNA.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE  1 (bases 190 to 292)
AUTHORS   Yamamoto,T., Tamura,T., Ryoji,M., Kaji,A., Yokota,T. and Takano,T.
TITLE      Sequence analysis of the heat-labile enterotoxin subunit B gene
            originating in human enterotoxigenic Escherichia coli
JOURNAL    J. Bacteriol. 152 (1), 506-509 (1982)
MEDLINE    83007048
AUTHORS   Yamamoto,T., Tamura,T., Yokota,T. and Takano,T.
TITLE      Overlapping genes in the heat-labile enterotoxin operon originating
            from Escherichia coli human strain
JOURNAL    Mol. Gen. Genet. 188 (2), 356-359 (1982)
MEDLINE    83114628
AUTHORS   Yamamoto,T. and Yokota,T.
TITLE      Sequence of heat-labile enterotoxin of Escherichia coli pathogenic
            for humans
JOURNAL    J. Bacteriol. 155 (2), 728-733 (1983)
MEDLINE    83265593
COMMENT    see also other loci beginning <ecoelt>.
FEATURES   location/Qualifiers
            source          1..774
                        /organism="Escherichia coli"
                        /strain="H10407"
                        /specific_host="Homo sapiens"
                        /db_xref="taxon:562"
                        <1..208
                        /codon_start=2
                        /transl_table=11
                        /product="enterotoxin A"
                        /protein_id="AA02981.1"
                        /db_xref="GI:1395121"
                        /translation="AMREPMTHHAPOGGDSRRITGPTCNBERQNLSTIYLKRYQS
                        KYKROTFSDYQSEVDIYNIRNEL"
                        195..197
                        /gene="tox"
                        195..579
                        /gene="tox"
                        205..267
                        /gene="tox"
                        205..579
                        /gene="tox"
                        /codon_start=1
                        /transl_table=11
                        /product="enterotoxin B"
                        /protein_id="AA02982.1"
                        /db_xref="GI:1395122"
                        /translation="MNKVKCYVFTALLSSICAYGAPGSITEICSEYRNPTQIYTTNDK
                        ILSYTESMAGKREMYITPFKSGATFOVEVPGSOHIDSOKKAIERMKDITLITVLTETK
                        IDKICWMNKTPTNSIAISMEN"

```

CDS

```

RBS
gene
sig_peptide
CDS

```

```

/gene="tox"
/gene="tox"
/gene="tox"
/gene="tox"
/gene="tox"
/codon_start=1
/transl_table=11
/product="enterotoxin B"
/protein_id="AA02982.1"
/db_xref="GI:1395122"
/translation="MNKVKCYVFTALLSSICAYGAPGSITEICSEYRNPTQIYTTNDK
ILSYTESMAGKREMYITPFKSGATFOVEVPGSOHIDSOKKAIERMKDITLITVLTETK
IDKICWMNKTPTNSIAISMEN"

```


mat_peptide 268 .576
 /gene="toxh"
 /product="enterotoxin B"
 BASE COUNT 286 a 142 c 153 g 193 t
 ORIGIN

alignment_scores:
 Quality: 24.00 Length: 24
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x ECOELT ..

Align seg 1/1 to: ECOELT from: 1 to: 774

214 TyrglnSerLysValLysArgGlnIlePheSerAspTyrGlnSerGluVa 230
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 125 TATCATATCAAAAGTTAAAGAGCAGATATTTCAGACTATCAGTCAGAGCT 174
 230 LAspIleTyrAsnArgIleArg 237
 ||||||||||||||||||||||||||||
 175 TGACATATATTAACAGATTCGG 196

seq_name: gb_ba:ECOPHLEGS

seq_documentation_block:

LOCUS ECOPHLEGS 900 bp DNA linear BCT 26-APR-1993
 DEFINITION Escherichia coli heat-labile enterotoxin gene, 5' cds.
 ACCESSION M61015
 VERSION M61015.1 GI:147190
 KEYWORDS heat-labile enterotoxin.
 SOURCE E.coli, DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE
 1 (bases 1 to 900)
 Trachman,J.D. and Maas,W.K.
 Characterization of the control region of the porcine heat-labile
 enterotoxin in Escherichia coli

Unpublished (1991)

JOURNAL
 FEATURES
 source Location/Qualifiers

1..900
 /organism="Escherichia coli"
 /strain="K-12"
 /db_xref="taxon:562"
 722..>900
 /product="heat-labile enterotoxin"
 781..>900
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin"
 /protein_id="AA24335.1"
 /db_xref="GI:147191"
 /translation="MKNIFFIFILLASPLVANGDRLYRADSRPDEIRSGSL"

BASE COUNT 227 a 213 c 181 g 279 t
 ORIGIN

alignment_scores:
 Quality: 22.00 Length: 22
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOPHLEGS ..

Align seg 1/1 to: ECOPHLEGS from: 1 to: 900

1 AsngLysArgLeuTyrArgAlaSerArgProProAspGluIleIlely 17
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 835 AATGGCGACAGATATATACCGTGTACTGTAGACCCCGAGATGAATAATA 884

17 sargSerGlyGlyLeu 22
 ||||||||||||||||
 885 ACGTTCGGAGAGTCTT 900

seq_name: gb_ba:VIBCTA1

seq_documentation_block:

LOCUS VIBCTA1 255 bp DNA linear BCT 26-APR-1993
 DEFINITION Vibrio cholerae cholera gene: subunit A alpha-chain 5' end.
 ACCESSION J01846
 VERSION J01846.1 GI:155156
 KEYWORDS cholera toxin.
 SEGMENT 1 of 2
 SOURCE Vibrio cholerae (strain El Tor 1621) DNA.
 ORGANISM Vibrio cholerae

REFERENCE
 1 (bases 1 to 255)
 Gennaro,M.L. and Greenaway,P.J.

Nucleotide sequences within the cholera toxin operon
 JOURNAL Nucleic Acids Res. 11, 3855-3861 (1983)
 MEDLINE 83246519
 COMMENT clean copy provided by authors.

FEATURES
 source Location/Qualifiers

1..255
 /organism="Vibrio cholerae"
 /db_xref="taxon:666"
 BASE COUNT 85 a 23 c 38 g 109 t
 ORIGIN

alignment_scores:
 Quality: 13.00 Length: 13
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTA1 ..

Align seg 1/1 to: VIBCTA1 from: 1 to: 255

5 LeuTyrArgAlaAspSerArgProProAspGluIleLys 17
 ||||||||||||||||||||||||||||
 199 TTATATCGGCGAGATCTAGACTCTGTAGTGAATAAAG 237

seq_name: gb_ba:VIBCTA2

seq_documentation_block:

LOCUS VIBCTA2 251 bp DNA linear BCT 26-APR-1993
 DEFINITION Vibrio cholerae cholera toxin gene: subunit A gamma-chain 3' end &
 subunit B 5' end.
 ACCESSION J01847
 VERSION J01847.1 GI:155157
 KEYWORDS cholera toxin.
 SEGMENT 2 of 2
 SOURCE Vibrio cholerae (strain El Tor 1621) DNA.
 ORGANISM Vibrio cholerae

REFERENCE
 1 (bases 1 to 251)
 Gennaro,M.L. and Greenaway,P.J.

Nucleotide sequences within the cholera toxin operon
 JOURNAL Nucleic Acids Res. 11, 3855-3861 (1983)
 MEDLINE 83246519

FEATURES
 source Location/Qualifiers

1..251
 /organism="Vibrio cholerae"
 /db_xref="taxon:666"
 BASE COUNT 96 a 38 c 38 g 79 t
 ORIGIN

alignment_scores:
 Quality: 12.00 Length: 12

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTA2 ..

Align seg 1/1 to: VIBCTA2 from: 1 to: 251

213 GluTyrGlnSerLysValLysArgGlnIlePheSer 224
|||||
61 GAATACCAATCTAAAGTTAAAGACAATATTTTCA 96

seq_name: gb_ba:VIBTOX

seq_documentation_block:

LOCUS VIBTOX 512 bp DNA linear BCT 26-APR-1993
DEFINITION Vibrio cholerae toxA and toxB genes for cholera enterotoxin

subunits A2 (gamma) and B.

ACCESSION K01170.1 GI:155296

VERSION enterotoxin.

KEYWORDS Vibrio cholerae (El Tor biotype strain 62746) DNA, clone pCVD002.

SOURCE Vibrio cholerae

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 512)

AUTHORS Lockman, H. and Kaper, J. B.

TITLE Nucleotide sequence analysis of the A2 and B subunits of Vibrio

cholerae enterotoxin

JOURNAL J. Biol. Chem. 258, 13722-13726 (1983)

MEDLINE 84061784

COMMENT The termination codon of subunit A2 overlaps the initiation codon of subunit B by 4 bp. The correct translation of the 2 subunits is the result of a shift in reading frames at this junction. Previous protein structural analysis has shown that the A2 subunit is bound to a precursor of the A subunit. Therefore, the 'met' at position 1 is not an initiator.

FEATURES
Location/Qualifiers
1..512
/organism="Vibrio cholerae"
/db_xref="taxon:666"

CDS
1..141
/note="enterotoxin subunit A2"
/codon_start=1
/transl_table=11
/protein_id="AA027572.1"
/db_xref="GI:155297"
/translation="MSNFCDEKTSGLGVKELDEYQSKYKRIQIFSGYQSDIDTHNRIRKD EL"

CDS
138..512
/note="enterotoxin subunit B"
/codon_start=1
/transl_table=11
/protein_id="AA027573.1"
/db_xref="GI:155298"
/translation="MIKLFQVFTVLLSSAYAHGTPONTDLCAEYHNTQIYTLNDK IFSYTESLAGKREMAITFEKSGTEQVEVPGSOHDSOKKAIERMKDTRLRIYLYLEAK VEKLCVMNKKTPHAIATISMN"

sig_peptide
141..200
/note="enterotoxin subunit B signal peptide"

mat_peptide
201..509
/product="enterotoxin subunit B mature peptide"

BASE COUNT 199 a 74 c 88 g 151 t

ORIGIN 188 bp upstream of NdeI site.

alignment_scores:

Quality: 12.00 Length: 12
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBTOX ..

Align seg 1/1 to: VIBTOX from: 1 to: 512

213 GluTyrGlnSerLysValLysArgGlnIlePheSer 224
|||||
55 GAATACCAATCTAAAGTTAAAGACAATATTTTCA 90

seq_name: gb_ba:ECOELTBP

seq_documentation_block:

LOCUS ECOELTBP 604 bp DNA linear BCT 16-APR-1996
DEFINITION E.coli (from pig) heat-labile enterotoxin subunit B gene (eltB), complete cds.

ACCESSION M17873.1 GI:145832

VERSION M17873.1

KEYWORDS eltB gene; enterotoxin.

SOURCE Escherichia coli DNA.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 407)

AUTHORS Dallas, W.S. and Falkow, S.

TITLE Amino acid sequence homology between cholera toxin and Escherichia coli heat-labile toxin

JOURNAL Nature 288 (5790), 499-501 (1980)

MEDLINE 81074965

REFERENCE 2 (bases 17 to 604)

AUTHORS Leong, J., Vinal, A.C. and Dallas, W.S.

TITLE Nucleotide sequence comparison between heat-labile toxin B-subunit

cistrons from Escherichia coli of human and porcine origin

JOURNAL Infect. Immun. 48 (1), 73-77 (1985)

MEDLINE 85156481

COMMENT [2] reviews [1].

A long inverted repeat which could be a termination signal is

located at positions 533-564.

FEATURES
Location/Qualifiers
1..604
/organism="Escherichia coli"
/specific_host="Sus scrofa"
/db_xref="taxon:562"

gene
33..407
/gene="eltB"

sig_peptide
33..95
/gene="eltB"

CDS
33..407
/note="precursor"
/codon_start=1
/transl_table=11
/product="enterotoxin"
/protein_id="AA098065.1"
/db_xref="GI:145833"
/translation="MNKVKCYLFTALSSLYAHKAPOTTTELCSERYRNTQIYTLNDK ILSYTESMAGKREMYITFEKSGTEQVEVPGSOHDSOKKAIERMKDTRLRIYLYLETK IDKLCVMNKKTPHAIATISMN"

mat_peptide
96..404
/gene="eltB"

old_sequence
114
/product="enterotoxin"

old_sequence
114
/gene="eltB"

old_sequence
223
/note="Nature 288, 499-501 (1980)"

old_sequence
223
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

old_sequence
362
/citation=[1]

old_sequence
362
/gene="eltB"

old_sequence
362
/note="Nature 288, 499-501 (1980)"

alignment_scores:
Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ECOELTBP ..

Align seg 1/1 to: ECOELTBP from: 1 to: 604

230 ValaspietTyrasnArgilearAspGluLeu 240

1 GTTGACATATATACAGATTCGGGATGAATTA 33

seq_name: gb_ba:ECOETOXHL

seq_documentation_block:

LOCUS ECOETOXHL 1262 bp DNA linear BCT 18-NOV-1994

DEFINITION Escherichia coli (strain 41) heat-labile enterotoxin type IIB

ACCESSION M28523.1 GI:576584

KEYWORDS LT-IIB gene; enterotoxin type IIB.

SOURCE Escherichia coli (strain 41) DNA.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1 to 1262)

AUTHORS Pickett,C.L., Twiddy,E.M., Coker,C. and Holmes,R.K.

TITLE Cloning, nucleotide sequence, and hybridization studies of the type

JOURNAL J. Bacteriol. 171 (9), 4945-4952 (1989)

COMMENT 89359131

FEATURES

source

-35_signal

gene

-10_signal

CDS

sig_peptide

mat_peptide

sig_peptide

mat_peptide

sig_peptide

mat_peptide

sig_peptide

mat_peptide

sig_peptide

mat_peptide

/protein_id="AAA53286.1"
/db_xref="GI:576586"
/translation="MSFKIITKAFVIMATLVVOAHAGASQFFRDNCRRTASLVEGV
ELTKYISDINNNDGMYVSSGCVWRISRAKDDPDVMTAEMRKIAAAVILSGMRVN
MCASPASSPNYIMALELEAF"
963..1259

/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"

BASE COUNT 369 a 195 c 285 g 413 t

ORIGIN

alignment_scores:

Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOETOXHL ..

Align seg 1/1 to: ECOETOXHL from: 1 to: 1262

38 AsnIleasnLeuTyrAspHisAlaArgGlyThr 48

278 AACATTAACCTATATGACATGACGAGTACT 310

seq_name: gb_pat:AR019034

seq_documentation_block:

LOCUS AR019034 42 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 14 from patent US 5783196.

ACCESSION AR019034

KEYWORDS AR019034.1 GI:3974148

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 42)

AUTHORS Noriega,F.R. and Levine,M.M.

TITLE Gua mutants of shigella spp. and vaccines containing the same

JOURNAL Patent: US 5783196-A 14 21-JUL-1998;

FEATURES

source

BASE COUNT 14 a 8 c 10 g 10 t

ORIGIN

alignment_scores:

Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AR019034 ..

Align seg 1/1 to: AR019034 from: 1 to: 42

192 ArgThrIleThrGlyAspThrCysAsnGlu 201

13 AGAACAAATTAACGAGTACTGTAAATGAG 42

/protein_id="AAA53286.1"
/db_xref="GI:576586"
/translation="MSFKIITKAFVIMATLVVOAHAGASQFFRDNCRRTASLVEGV
ELTKYISDINNNDGMYVSSGCVWRISRAKDDPDVMTAEMRKIAAAVILSGMRVN
MCASPASSPNYIMALELEAF"
963..1259

/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"

BASE COUNT 369 a 195 c 285 g 413 t

ORIGIN

alignment_scores:

Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOETOXHL ..

Align seg 1/1 to: ECOETOXHL from: 1 to: 1262

38 AsnIleasnLeuTyrAspHisAlaArgGlyThr 48

278 AACATTAACCTATATGACATGACGAGTACT 310

seq_name: gb_pat:AR019034

seq_documentation_block:

LOCUS AR019034 42 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 14 from patent US 5783196.

ACCESSION AR019034

KEYWORDS AR019034.1 GI:3974148

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 42)

AUTHORS Noriega,F.R. and Levine,M.M.

TITLE Gua mutants of shigella spp. and vaccines containing the same

JOURNAL Patent: US 5783196-A 14 21-JUL-1998;

FEATURES

source

BASE COUNT 14 a 8 c 10 g 10 t

ORIGIN

alignment_scores:

Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AR019034 ..

Align seg 1/1 to: AR019034 from: 1 to: 42

192 ArgThrIleThrGlyAspThrCysAsnGlu 201

13 AGAACAAATTAACGAGTACTGTAAATGAG 42

seq_name: gb_pat:A69697

seq_documentation_block:

LOCUS A69697 2724 bp DNA linear PAT 07-MAY-1999

DEFINITION Sequence 15 from Patent WO9807864.

ACCESSION A69697

VERSION A69697.1 GI:4774310

KEYWORDS unidentified.

SOURCE unidentified.

ORGANISM unidentified


```

REFERENCE 1 (bases 1 to 2724)
AUTHORS Foster K.A., Quinn, C.P. and Shone, C.C.
TITLE RECOMBINANT TOXIN FRAGMENTS
JOURNAL Patent: WO 9807864-A 15 26-FEB-1998;
FOSTER KEITH ALAN (GB)
FEATURES
source 1..2724
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
1..2721
mat_peptide 1..2721
/product="unnamed"
1..2724
CDS
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAB42522.1"
/db_xref="GI:4774311"
/translacion="MQFVNKQFNYKDPVNGVDIAYIKIIPNAGOMQPVKAKIHNKIIV
IPERDTFNPEEGDLNPPPEAKQVSVYSDSYLSTDNKKNYLGKVTLEPRITSD
LGRMLTSIVRGIPWGGSTIDTELKVIDTNCINVIOPGSGVRSSEINLVITGPSADI
IOFECKSRGHEVLANTRNGYSTOYIRSPDTRFPEESLEVDTPNPLGACFATDPA
VTLAHELHAGHRLGIAINPVRKVTNMAIYEWGSELEVEELRTEGHDAKTDS
LOENEFRLYYNKKFEDIASTLNKAKSIYGTASIQMKVNEKYLLESDTGKFSVD
KLKFDKLYKMLTEIYEDNPFVFEFVNLNKRITLNFDAVKINIPKVVYIYDGFNL
RTNLANPNGONTFELNNMNFETLKNFTGLCFEYKILCYRGITTSKSLDGYNKIE
GRCDGALNDICIKVNMNDFEESSEDNFTNDLNKGEITSDNINIAEENISLIDIOO
YLTFRPNENPENISIEUNSSDIQLELMPRIEERPNKKEIENDKYTFHFLRAQEF
EHGKSRILINSVNEALNPSRVYTFEESDVKYKAKNKAEMFGLWEQVLVDFDE
LTSVSTPKIADITIIIPYIGPALNIGMLYKDDVGAIFSGAVILFEIPEIAIPV
LTFALVSYIANKVLTVOTIDNLSKRNEKMDVYKVIYTNMLAVNVOIDILRRKM
EALENQASMTKAIINYQYNOYEERKNINENIDLSKLNSEINKAMINIKPLNOC
SVSYLNMASIPYGVKLEDEPDLAKLYIYDNRGLIGOVDRKDKVNTLSIDI
PQULSKYVNDNRLLSTFTEYIKSRQSKYKRIEFGYQSDIDTNRINDEL"
BASE COUNT 974 a 418 c 509 g 823 t
ORIGIN
alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x A69697 ..
Align seg 1/1 to: A69697 from: 1 to: 2724
215 GlnSerlyVallyArgGlnIlePheSer 224
|||||
2644 CAATCTAAAGTTAAAGACAATATTTC 2673
seq_name: gb_pat:BD009884
seq_documentation_block:
LOCUS BD009884 2724 bp DNA linear PAT 31-JAN-2002
DEFINITION Recombinant toxin fragments.
ACCESSION BD009884
BD009884.1 GI:18638257
KEYWORDS JP 2001502890-A/8.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2724)
AUTHORS Shone, C.C., Quinn, C.P. and Foster, K.A.
TITLE Recombinant toxin fragments
JOURNAL Patent: JP 2001502890-A 8 06-MAR-2001;
MICROBIOLOGICAL RESEARCH AUTHORITY CAMR, THE SPEYWOOD LABORATORY LTD
OS Unidentified
COMMENT
PD JP 2001502890-A/8
PN 06-MAR-2001
PF 22-AUG-1997 JP 1998510524
PR 23-AUG-1996 GB 9617671.4.13-DEC-1996 GB 9625996.5 PI
CLIFFORD CHARLES SHONE, CONRAD PADRAIG QUINN, KEITH ALAN FOSTER PC

```

```

C12N15/31.C12N1/21.C12P21/02.C07K14/33.A61K38/16.A61K39/08 CC
Strandedness: Double:
CC Topology: Linear:
FH Key Location/Qualifiers
FT CDS 1..2724.
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 974 a 418 c 509 g 823 t
ORIGIN
alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x BD009884 ..
Align seg 1/1 to: BD009884 from: 1 to: 2724
215 GlnSerlyVallyArgGlnIlePheSer 224
|||||
2644 CAATCTAAAGTTAAAGACAATATTTC 2673
seq_name: gb_htg:AC068893
seq_documentation_block:
LOCUS AC068893 154611 bp DNA linear HTG 30-AUG-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-283C16, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC068893
AC068893.4 GI:13811856
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 154611)
AUTHORS Smith, D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
2 (bases 1 to 154611)
REFERENCE 2 (bases 1 to 154611)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On Apr 27, 2001 this sequence version replaced gi:11545931.
COMMENT
----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: hg273
----- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 167668 bases at least Q40
Consensus quality: 171341 bases at least Q30
Consensus quality: 173207 bases at least Q20
Insert size: 177782; sum-of-ctnigs
Quality coverage: 6.1x in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```



```

* 6875 6974: gap of unknown length
* 6975 8613: contig of 1639 bp in length
* 8614 8714: gap of unknown length
* 8714 9769: contig of 1056 bp in length
* 9770 9869: gap of unknown length
* 9870 11095: contig of 1226 bp in length
* 11096 11195: gap of unknown length
* 11379 13378: contig of 2183 bp in length
* 13379 13478: gap of unknown length
* 13479 17449: contig of 3971 bp in length
* 17450 17549: gap of unknown length
* 20362 20461: contig of 812 bp in length
* 20462 24773: contig of 4312 bp in length
* 24774 28965: gap of unknown length
* 28966 29065: gap of unknown length
* 29066 34516: contig of 5451 bp in length
* 34517 42471: contig of 7855 bp in length
* 42472 42571: gap of unknown length
* 42572 47944: contig of 5373 bp in length
* 47945 48044: gap of unknown length
* 48045 54711: contig of 6667 bp in length
* 54712 54811: gap of unknown length
* 54812 63169: contig of 8358 bp in length
* 63170 74889: gap of unknown length
* 74890 74990: gap of unknown length
* 74990 87514: contig of 12525 bp in length
* 87515 87614: gap of unknown length
* 87615 95455: contig of 7841 bp in length
* 95456 95555: gap of unknown length
* 95556 125405: contig of 29850 bp in length
* 125406 125505: gap of unknown length
* 125506 149761: contig of 24156 bp in length
* 149762 166872: gap of unknown length
* 166873 17111: contig of 17111 bp in length.

```

FEATURES

source

```

1..166872
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-310A17"
/clone_1bp="RPCT-11"
1..1139
/feature="assembly_name:Contig1"
1240..2468
/feature="assembly_name:Contig9"
2569..3970
/feature="assembly_name:Contig11"
4071..5350
/feature="assembly_name:Contig13"
5651..6874
/feature="assembly_name:Contig14"
6975..8613
/feature="assembly_name:Contig15"
8714..9769
/feature="assembly_name:Contig17"
9870..11095
/feature="assembly_name:Contig19"
11196..13378
/feature="assembly_name:Contig20"
clone_end:SP6
13479..17449
/feature="assembly_name:Contig21"
17550..20361
/feature="assembly_name:Contig22"
20462..24773
/feature="assembly_name:Contig23"
24874..28965
/feature="assembly_name:Contig24"
29066..34516

```

```

/feature="assembly_name:Contig25"
clone_end:T7"
34617..42471
/feature="assembly_name:Contig26"
42572..47944
/feature="assembly_name:Contig27"
48045..54711
/feature="assembly_name:Contig28"
54812..63169
/feature="assembly_name:Contig29"
63270..74889
/feature="assembly_name:Contig30"
74990..87514
/feature="assembly_name:Contig31"
87615..95455
/feature="assembly_name:Contig32"
95556..125405
/feature="assembly_name:Contig33"
125506..149661
/feature="assembly_name:Contig34"
149762..166872
/feature="assembly_name:Contig35"

```

BASE COUNT 47731 a 34417 c 33394 g 49026 t 2304 others

ORIGIN

alignment_scores:

Quality:	10.00	Length:	10
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AC021033 ..

Align seg 1/1 to: AC021033 from: 1 to: 166872

62 ThrsEulerSeqArgSerAlaHisLeu 71

141955 ACTTCTTTGTCCCTTAGGAGTGCACACCTC 141984

```

seq_name: gb_pr:AC073367
seq_documentation_block:
LOCUS AC073367 168918 bp DNA linear PRI 25-OCT-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-197M22, complete sequence.
ACCESSION AC073367
VERSION AC073367.10 GI:16418121
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Oct 25, 2001 this sequence version replaced gi:15553107.
FEATURES
source
1..168918
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"

```


BASE COUNT 49033 a 34070 c 34897 g 50918 t
ORIGIN /clone="RP11-197M22"
 /clone_11b="RPC1-11"

alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AC073367/rev ..

Align seg 1/1 to reverse of: AC073367 from: 1 to: 168918

62 ThrSerLeuSerLeuArgSerAlaHisLeu 71
|||||
104393 ACTTCTTGTCTCCCTAAGAGTGACACACCTTC 104364

OM of: US-09-528-682-1 to: N.Geneseq_032802.* out_format : pfs
Date: Jun 18, 2002 10:55 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame1.p2n.model -DEV=xlp  
-O=/cgn2_1/USPTO_spool/US09528682/runat_18062002_082606_8842/app.query.fasta_1.299  
-B=N.Geneseq_032802 -QFMT=fastap -SUFFIX=olip2n.rng  
-LOOPEXT=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPEXT=0.000  
-LOOPEXT=0.000 -GAPEXT=4.500 -MINMATCH=0.050 -XGAPEXT=60.000  
-XGAPEXT=60.000 -GAPEXT=7.000 -XGAPEXT=60.000  
-XGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=0150 -TRANS=human40.cdi -LIST=1000 -DOCLIGN=200  
-THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09528682 -CGNL_1.563 -NCPU=6 -ICPU=3 -LONGIOS  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPRX -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-528-682-1  
Database: N.Geneseq_032802.*  
Query length: 240  
Database length: 1736436  
Database length: 858457221  
Search time (sec): 270.910000
```

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:AAAS0205 +		240.00	4287.77	1.8e-230	77
/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAO23864 +		208.00	3710.84	1.4e-198	77
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1106 +		184.00	3282.99	1.6e-174	78
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1544 +		184.00	3282.99	1.6e-174	78
/SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:AAAS0206 +		179.00	3193.32	1.6e-169	77
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAS01506 +		128.00	2273.54	2.7e-167	77
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1147 +		125.00	2224.92	1.5e-115	78
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1545 +		116.00	2063.94	1.5e-106	78
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51314 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51315 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51316 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51317 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51318 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51319 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51320 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51321 +		86.00	1525.37	1.3e-76	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51322 +		84.00	1489.49	1.3e-74	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51323 +		84.00	1489.49	1.3e-74	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51324 +		84.00	1489.49	1.3e-74	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51325 +		81.00	1435.61	1.3e-72	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51326 +		74.00	1310.07	1.3e-64	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51327 +		74.00	1310.07	1.3e-64	71
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51328 +		30.00	520.00	1.3e-20	77
/SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:AAO30099 +		30.00	519.83	1.3e-20	79
/SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAO74431 +		26.00	454.00	6.1e-17	5
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51328 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51329 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51330 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51331 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51332 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51333 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51334 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAO46318 +		26.00	448.74	1.2e-16	72
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAO51157 +		26.00	448.23	1.3e-16	74

/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1158 +	26.00	448.23	1.3e-16
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1159 +	26.00	448.23	1.3e-16
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1160 +	26.00	448.23	1.3e-16
/SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:AAO30169 +	26.00	443.43	1.8e-16
/SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAO26244 +	26.00	441.26	3.1e-16
/SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:AAO12710 +	25.00	430.26	1.3e-15
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAO23061 +	25.00	430.26	1.3e-15
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAO23062 +	25.00	430.26	1.3e-15
/SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAO51327 +	24.00	412.86	1.2e-14
/SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAO31263 +	22.00	376.46	1.3e-12
/SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAO33063 +	15.00	271.27	9.2e-07
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAO22789 +	15.00	271.27	9.2e-07
/SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAO22790 +	15.00	271.27	9.2e-07
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1117 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1118 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1119 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1120 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1121 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1122 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1123 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1124 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1125 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1126 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1127 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1128 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1129 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1130 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1131 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1132 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1133 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1134 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1135 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1136 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1137 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1138 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1139 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1140 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1141 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1142 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1143 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1144 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1145 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1146 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1147 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1148 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1149 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1150 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1151 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1152 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1153 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1154 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1155 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1156 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1157 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1158 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1159 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1160 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1161 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1162 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1163 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1164 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1165 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1166 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1167 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1168 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1169 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1170 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1171 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1172 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1173 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1174 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1175 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1176 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1177 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1178 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1179 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1180 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1181 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1182 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1183 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1184 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1185 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1186 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1187 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1188 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1189 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1190 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1191 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1192 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1193 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1194 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1195 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1196 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1197 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1198 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1199 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1200 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1201 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1202 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1203 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1204 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1205 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1206 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1207 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1208 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1209 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1210 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1211 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1212 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1213 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1214 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1215 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1216 +	13.00	236.23	8.2e-05
/SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS1217 +			

[illegible]

/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF59601 +	7.00	107.44	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT74890 -	7.00	102.46	2.3e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH167447 -	7.00	107.36	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV42931 -	7.00	102.46	2.3e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH67353 -	7.00	107.32	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA62681 -	7.00	102.39	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT86879 +	7.00	107.32	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL04193 +	7.00	102.32	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV21397 +	7.00	107.32	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS74362 -	7.00	102.25	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV63198 +	7.00	107.32	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS73756 -	7.00	102.23	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39909 +	7.00	107.13	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS71276 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39907 +	7.00	107.02	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS71372 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39904 -	7.00	107.01	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF94435 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA91565 -	7.00	107.01	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH82780 -	7.00	102.09	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA93596 -	7.00	107.01	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH82780 -	7.00	102.09	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA67186 +	7.00	106.98	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAH87749 -	7.00	102.06	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA67186 +	7.00	106.98	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF64031 -	7.00	102.06	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAX39906 +	7.00	106.93	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84728 -	7.00	102.06	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAX39902 +	7.00	106.81	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12229 +	7.00	101.86	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH04625 -	7.00	106.75	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC91212 -	7.00	101.86	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS74361 -	7.00	106.75	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS91196 +	7.00	101.84	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC54189 -	7.00	106.62	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12230 +	7.00	101.83	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC52617 -	7.00	106.60	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12231 +	7.00	101.83	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA13985 +	7.00	106.40	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12231 +	7.00	101.80	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC87642 +	7.00	106.38	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAF67124 +	7.00	101.80	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA167448 -	7.00	106.36	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AAO54134 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA02641 -	7.00	106.22	1.4e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT44455 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAA02641 -	7.00	106.12	1.4e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59657 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC38786 -	7.00	106.13	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO92524 -	7.00	101.76	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC38786 -	7.00	106.13	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC54400 +	7.00	101.43	2.7e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS25967 +	7.00	106.11	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS68183 -	7.00	101.30	2.7e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT57067 -	7.00	106.05	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAZ65344 +	7.00	101.30	2.7e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT57067 -	7.00	106.02	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84472 -	7.00	100.99	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS72626 -	7.00	105.63	1.6e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF55125 +	7.00	100.97	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX20025 -	7.00	105.21	1.6e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82129 +	7.00	100.96	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO99460 -	7.00	105.09	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82129 +	7.00	100.93	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT29627 -	7.00	104.95	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV20862 +	7.00	100.93	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAT01475 -	7.00	104.95	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA46136 +	7.00	100.93	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC35915 -	7.00	104.86	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA47422 -	7.00	100.85	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS54152 +	7.00	104.83	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84479 +	7.00	100.84	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH73931 -	7.00	104.83	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL07521 +	7.00	100.68	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH02888 -	7.00	104.63	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAF09705 -	7.00	100.63	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC47426 +	7.00	104.58	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAF09066 -	7.00	100.63	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC373980 +	7.00	104.52	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAS93383 -	7.00	100.63	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC373980 +	7.00	104.52	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS94796 +	7.00	100.51	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC49323 -	7.00	104.45	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC47954 +	7.00	100.50	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA96224 -	7.00	104.45	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL26737 +	7.00	100.47	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS33697 -	7.00	104.44	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA33152 +	7.00	100.47	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX20024 -	7.00	104.44	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAO79704 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS88441 +	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV09065 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77742 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX99382 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77743 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX99382 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77745 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD11111 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77745 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD26582 +	7.00	99.79	3.2e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK87870 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL24490 -	7.00	99.79	3.2e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK87870 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO15170 -	7.00	99.60	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77744 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV41879 -	7.00	99.60	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK78705 -	7.00	104.43	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO38091 -	7.00	99.52	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS53102 +	7.00	104.40	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS159387 -	7.00	99.40	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF6849 +	7.00	104.28	1.9e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78706 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF6849 +	7.00	104.28	1.9e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX81526 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF64445 -	7.00	104.12	1.9e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX81887 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC33336 +	7.00	104.05	1.9e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV15527 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS79299 +	7.00	103.69	2.0e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV15519 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS79299 +	7.00	103.69	2.0e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX83888 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB14359 -	7.00	103.65	2.0e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX83888 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB14359 -	7.00	103.65	2.0e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX83888 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF21703 +	7.00	103.45	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC46027 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF67123 +	7.00	103.38	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT44581 -	7.00	99.25	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAZ62626 -	7.00	103.28	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD06040 +	7.00	99.19	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL12897 -	7.00	103.16	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS44736 -	7.00	99.11	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ06250 -	7.00	103.13	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS43070 -	7.00	99.10	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB14359 -	7.00	103.12	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS43070 -	7.00	99.07	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB07752 +	7.00	103.10	2.1e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS43070 -	7.00	99.05	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT26989 -	7.00	103.06	2.2e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI12557 +	7.00	99.04	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS98184 -	7.00	103.00	2.2e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH16894 -	7.00	98.93	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS98184 -	7.00	102.95	2.2e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA35042 -	7.00	98.97	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC77954 -	7.00	102.71	2.3e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT95234 -	7.00	98.97	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS71553 -	7.00	102.59	2.3e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF99234 -	7.00	98.97	3.6e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK94170 -	7.00	102.56	2.3e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF99234 -	7.00	98.97	3.6e+03	7

/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AMS41961-	7.00	90.54	1.1e+04	13	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21613+	7.00	68.66	1.7e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AMS41962+	7.00	90.54	1.1e+04	8	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH51614+	7.00	66.40	2.3e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AMS45133+	7.00	89.90	1.2e+04	8	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH61689+	7.00	64.57	2.9e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABLO6513+	7.00	89.79	1.2e+04	9	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH66534+	7.00	64.47	2.9e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK86614-	7.00	89.34	1.3e+04	9	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21554-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:AAK86614-	7.00	89.34	1.3e+04	9	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21607-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:ABLO2636-	7.00	89.32	1.3e+04	9	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21609-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH63692+	7.00	88.73	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21610-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH62676+	7.00	88.65	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21611+	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:ABH617191+	7.00	88.64	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21612+	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2002.DAT:ABJ320319+	7.00	88.56	1.4e+04	10	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE21612-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH79031+	7.00	88.39	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH41223+	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK79984-	7.00	88.35	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH41226-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH6162703+	7.00	88.39	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH64966-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK79983+	7.00	88.33	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH66530-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH6162708+	7.00	88.33	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH66531-	7.00	63.58	3.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH545479+	7.00	88.33	1.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAE22303+	7.00	59.59	5.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2002.DAT:AAH33984+	7.00	88.12	1.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH81489+	7.00	57.34	6.8e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH530105-	7.00	88.00	1.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH81489-	7.00	57.34	6.8e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH12965+	7.00	87.88	1.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH91930-	7.00	54.58	9.3e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000A.DAT:AAH530100+	7.00	87.79	1.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH91930-	7.00	53.47	1.0e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000A.DAT:AAH81724+	7.00	87.69	1.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH952740+	7.00	53.14	1.1e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH545671+	7.00	87.52	1.6e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH96733+	7.00	53.14	1.1e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:AAH595370+	7.00	87.42	1.6e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH99683+	7.00	45.45	2.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH13275+	7.00	87.08	1.7e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH99682+	7.00	45.44	2.2e+00	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABLO2388-	7.00	86.88	1.7e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAQ32800-	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAQ40862+	7.00	86.82	1.7e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAQ48097+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH15410-	7.00	86.29	1.9e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAQ48098+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH850593+	7.00	85.92	1.9e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAQ50101-	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH059533+	7.00	85.90	1.9e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH51698-	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH059534+	7.00	85.89	1.9e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH58891+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH059535+	7.00	85.89	1.9e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH58942+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK73340-	7.00	85.88	2.0e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH58949+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK73337-	7.00	85.88	2.0e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH59785+	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK73338-	7.00	85.88	2.0e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH59780-	6.00	115.96	412.03	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK73363-	7.00	85.88	2.0e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAQ21399-	6.00	115.60	431.91	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH12468-	7.00	85.37	2.1e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAQ26798+	6.00	115.60	431.91	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH18110-	7.00	85.37	2.1e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAQ27358+	6.00	115.60	431.91	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH18910-	7.00	85.37	2.1e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAQ32803+	6.00	115.25	451.70	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH20510+	7.00	85.08	2.2e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAH48101+	6.00	115.25	451.70	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH199469+	7.00	84.89	2.2e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAH48101+	6.00	115.25	451.70	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH164114+	7.00	84.89	2.2e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH58953+	6.00	115.25	451.70	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH494200-	7.00	84.86	2.2e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH58959+	6.00	115.25	451.70	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH76759-	7.00	84.85	2.2e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH79789+	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAK66643+	7.00	84.81	2.4e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH505100-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH16796-	7.00	83.98	2.5e+04	1	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH505100-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH132829+	7.00	83.88	2.5e+04	20	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH47548-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH043024+	7.00	83.78	2.6e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH47548-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH043029+	7.00	83.65	2.6e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH16278-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA1996.DAT:AAH11658-	7.00	83.65	2.6e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH12729-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH05882+	7.00	83.24	2.7e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH12729-	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2002.DAT:ABH05882+	7.00	83.24	2.7e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH258546+	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:ABH05882+	7.00	83.24	2.7e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH258546+	6.00	114.92	471.41	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:ABH05882+	7.00	83.24	2.7e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAH046053+	6.00	114.29	510.63	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH18618-	7.00	82.77	2.9e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH046053+	6.00	114.29	510.63	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH19608+	7.00	82.66	3.0e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH21358+	6.00	114.00	530.14	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:AAH59608+	7.00	82.61	3.0e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH20286+	6.00	114.00	530.14	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH11177-	7.00	82.25	3.1e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH384459-	6.00	114.00	530.14	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH70447+	7.00	81.91	3.3e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAH513125+	6.00	113.72	549.58	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH91264-	7.00	81.91	3.3e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA1997.DAT:AAH513125+	6.00	113.72	549.58	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH91264-	7.00	81.29	3.5e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH86531+	6.00	113.72	549.58	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH332517+	7.00	81.28	3.5e+04	2	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH86531+	6.00	113.72	549.58	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH13678-	7.00	80.93	3.7e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH465501-	6.00	113.72	549.58	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH13678-	7.00	80.93	3.7e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA1998.DAT:AAH03367+	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH13678-	7.00	80.60	3.8e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH12954-	7.00	80.54	3.9e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH51294-	7.00	80.49	3.9e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH501284+	7.00	79.49	4.4e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH501284+	7.00	79.49	4.4e+04	3	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH501284+	7.00	78.68	4.9e+04	4	/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH260782-	6.00	113.45	568.96	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH615122+	7.00	78.34	5.1e+04	4	/SID5/gcgdata/genseq/genseq-emb1/NA1999.DAT:AAH009440-	6.00	112.94	607.55	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH75080+	7.00	78.29	5.1e+04	4	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH20875-	6.00	112.94	607.55	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001B.DAT:ABH16830+	7.00	77.97	5.3e+04	4	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH20875-	6.00	112.69	626.76	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH67426+	7.00	76.80	6.2e+04	5	/SID5/gcgdata/genseq/genseq-emb1/NA1993.DAT:AAH053628+	6.00	112.69	626.76	13
/SID5/gcgdata/genseq/genseq-emb1/NA2000.DAT:AAH584712+	7.00	76.35	6.5e+04	5	/SID5/gcgdata/genseq/genseq-emb1/NA1994.DAT:AAH063939-	6.00	112.69	626.76	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH834782+	7.00	76.01	6.9e+04	6	/SID5/gcgdata/genseq/genseq-emb1/NA1997.DAT:AAH09673-	6.00	112.46	645.92	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH834782+	7.00	76.01	6.9e+04	6	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH09673-	6.00	112.46	645.92	13
/SID5/gcgdata/genseq/genseq-emb1/NA2001A.DAT:AAH834782+	7.00	76.01	6.9e+04	6	/SID5/gcgdata/genseq/genseq-emb1/NA1992.DAT:AAH09673-	6.00	112.46	645.92	13

/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AB197592-	6.00	112.01	684.10	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK57505-	6.00	103.52	2.0e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV29249+	6.00	111.80	703.11	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK61599+	6.00	103.52	2.0e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB3283+	6.00	111.80	703.11	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAK61418+	6.00	103.52	2.1e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK27561+	6.00	111.59	722.08	35	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB6398+	6.00	102.95	2.2e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAG9697+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAB9347+	6.00	102.77	2.2e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAG97657+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAB9362+	6.00	102.77	2.2e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAG2618+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAB8649+	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB5000-	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7747+	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB51126+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7751+	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB48342+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7752-	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB48342-	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7755-	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB48402+	6.00	110.63	816.29	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7752-	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAB6483+	6.00	110.46	835.01	41	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT7746+	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAAB8294+	6.00	109.79	909.53	42	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT8114-	6.00	102.71	2.3e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAB78295+	6.00	109.79	909.53	42	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT8123-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB67727+	6.00	109.48	946.59	43	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:ABAT8293-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB67727-	6.00	109.48	946.59	43	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK89146-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB69171-	6.00	109.48	946.59	47	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK89218-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAB88599-	6.00	109.48	946.59	47	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK89375-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAB88599-	6.00	109.48	946.59	47	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK89375-	6.00	102.31	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB4801+	6.00	109.04	965.07	48	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAI54741-	6.00	102.25	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAB40202-	6.00	109.04	1.0e+03	50	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAK12311+	6.00	102.25	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34267-	6.00	109.04	1.0e+03	50	/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK16517+	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	109.04	1.0e+03	50	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK37497-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34267+	6.00	109.04	1.0e+03	50	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK37497-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34267-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAAF63765+	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765+	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4e+03	
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF63765-	6.00	108.89	1.0e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK281942-	6.00	102.19	2.4	

/SID5_gcgdata/geneseq/geneseq-emb1/NA1993.DAT.AA060538 +	6.00	99.98	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK24713 -	6.00	98.43	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC01599 +	6.00	99.94	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK50713 -	6.00	98.43	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC02706 +	6.00	99.90	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK127726 -	6.00	98.43	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC15089 +	6.00	99.90	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK56693 -	6.00	98.43	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF68390 -	6.00	99.90	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAT70061 -	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAL22359 -	6.00	99.86	3.2e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAT700126 -	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS50048 +	6.00	99.82	3.3e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA55370 -	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS50058 +	6.00	99.82	3.3e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA55370 -	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK40062 +	6.00	99.63	3.3e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAS47176 +	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK24001 +	6.00	99.63	3.3e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB85499 +	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAC127154 +	6.00	99.63	3.3e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF17746 +	6.00	98.40	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC52251 -	6.00	99.48	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC12385 +	6.00	98.36	3.9e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC05630 -	6.00	99.44	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAV35353 +	6.00	98.30	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA31618 +	6.00	99.44	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAS34564 +	6.00	98.30	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA31715 -	6.00	99.44	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAV75489 +	6.00	98.27	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB647095 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS49279 +	6.00	98.24	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB64977 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS450480 +	6.00	98.24	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK39140 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAV505624 +	6.00	98.20	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK19949 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC14438 +	6.00	98.17	4.0e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AA145617 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB76445 +	6.00	98.08	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AA105631 +	6.00	99.40	3.4e+03	17	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB25080 +	6.00	98.08	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC145268 -	6.00	99.37	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AA157125 +	6.00	98.08	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC44265 -	6.00	99.37	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAH72234 -	6.00	98.08	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC19356 +	6.00	99.29	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAH72234 -	6.00	98.08	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAA78976 +	6.00	99.25	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA39213 -	6.00	98.05	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA458161 +	6.00	99.22	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA39213 -	6.00	98.05	4.1e+03	21
/SID5_gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC30308 +	6.00	99.18	3.5e+03	19	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK79758 +	6.00	98.05	4.1e+03	21

/SIDS5/gcgdata/geneseq/geneseg-emb1/NA199. DAT:AAV90316 - 6.00 97.37 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA199. DAT:AAV9341 + 6.00 97.37 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAH71170 - 6.00 97.37 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAH81931 - 6.00 97.37 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2000. DAT:AAK08770 + 6.00 97.34 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAK61074 - 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA1995. DAT:AAV21162 - 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAV242961 + 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:ABA69945 + 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAK36786 + 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAK18155 + 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAK44046 + 6.00 97.31 4.5e+03 25
/SIDS5/gcgdata/geneseq/geneseg-emb1/NA2001A. DAT:AAK55337 - 6.00 97.31 4.5e+03 25

seq_name: /SIDS5/gcgdata/geneseq/geneseg-emb1/NA1985. DAT:AAV50205

seq_documentation_block:
ID AAV50205 standard; DNA: 777 BP.

XX AAV50205;
XX
DT 30-OCT-1991 (first entry)
XX
DE Sequence of the pig scours heat labile toxin (LT) LTA gene.
XX
KW Pig scours vaccine; toxin; diarrhoea; ss.
OS E.coli NCIB 11932.
XX
FH Key Location/Qualifiers
FT CDS 1..777
/*lag- a
XX
PN EPI45486-A.
XX
PD 19-JUN-1985.
XX
PE 12-DEC-1984; 84EP-0308620.
XX
PR 12-DEC-1983; 83GB-0033131.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Hayes MV, Harford S, Ross GW;
XX
DR WPI: 1985-148358/25.
DR P-PSDB; AAP50190.
XX
PT New toxoid as inactivated form of toxin for use in vaccines - is
PT obtd. from organism transformed by gene
XX
PS Disclosure; Fig 1; 61pp; English.
XX
CC AAV50205 is the gene sequence of the natural LTA gene. The LTA gene of
CC the site directed mutant SDM1 (see AAV50206) is inactive. The
CC inventors claim a vaccine prepn. active against pig scours which
CC contains an inactivated LTA component, together with additional K88
CC antigens opt. with whole cells comprising the antigens or contg. the
CC inactivated LTA.
XX
SQ Sequence 777 BP; 255 A; 136 C; 164 G; 222 T; 0 other;

alignment_scores:
Quality: 240.00 Length: 240
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAV50205 ..

Align seg 1/1 to: AAV50205 from: 1 to: 777

1 AsnGlyAspArgLeuTyrFargAlaAspSerArgProProaspGluIleTyr 17
|||||
55 AATGGCGACAGATATATACCGTGTGACGCTGACCCCGACATGAATATAA 104
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
|||||
105 ACGTTCGCGAGGCTGTATAGCCCAAGGCGCAATAAAGACTACTTCATAGAG 154
34 LTTThlMetAsnIleAsnLeuTyrAspHisAlaArgGlyThlGluThr 50
|||||
155 GACTCAATATATATATATTTATGTATGATCAGCGAGAGAACCAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThlSerLeuSerLeu 67
205 GCGTTTGTACAGATATGATGACGAGATATGTTCCACTTCTTCTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThlTyr 84
255 AAGTCTCCTACTTACGACGACAGCTATATATATATATATATATATATAT 304
84 YrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTTATATACGACGACCAATATATGTTATATGTTATATATAT 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
355 TTAGCGGTATACAGCCCTCACCACCATATGACAGAGGTTCTCGCGTTAG 404
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
405 TGGATATACCATATCTCAGATATATGATGATGATGATGATGATGATGAT 454
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
455 TGATGATGACAGATATATATGATGATGATGATGATGATGATGATGATGAT 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
505 AGAATCTGAAATATATGCTCCGCGAGAGATGTTACAGATATGACGATTT 554
167 eProProAspHisGlnAlaTyrPArgGluGluProTyrPheHisAlaP 184
555 CCCACGGATCACCAAGCTTGAGAGAGAACCCCTGGATTCATATGATCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
605 CACAAGGTGTGGAATTCATCAAGAACATACACAGGTGATCTGTATAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
655 GAGGAGACCCAGAAATCTGACCAATATATCTCAGGCAATATCAATCAA 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
705 AGTTAAGAGGCAAGATATTTTACACTATACAGTACAGAGGTGATATATA 754
234 snArgIleArgAspGluLeu 240
755 ACAGATTCGGGATCAATTA 774

seq_name: /SIDS5/gcgdata/geneseq/geneseg-emb1/NA1992. DAT:AAQ23864

seq_documentation_block:
ID AAQ23864 standard; DNA: 1148 BP.

XX AAQ23864;
XX
DT 08-OCT-1992 (first entry)
XX
DE Thermolabile toxin (LT) gene.
XX
KW LTH; ST1a; ST1b; thermolabile; heat-stable; E.coli; T4DNA ligase;
probe; detection; ss.


```

XX OS Escherichia coli.
XX PN JP04079898-A.
XX PD 13-MAR-1992.
XX PF 23-JUL-1990; 90JP-0194208.
XX PR 23-JUL-1990; 90JP-0194208.
XX PA (KITA ) KITASATO RES INST.
XX DR WPI; 1992-137930/17.
XX PT DNA and RNA probe - for simultaneous detection of Vibrio cholerae
XX PF and entero-toxicogenic E. coli
XX PS
XX PS Claim 1; Page 1; 15pp; Japanese.
XX CC The thermostable toxin (Lth) gene sequence is derived from
XX CC enterotoxigenic E. coli. An Lth-ST1a-ST1b probe, wherein all
XX CC all or part of the sequences of the three toxin genes represented
XX CC in AA023864-66 are ligated to each other by T4DNA ligase, may be used
XX CC for the detection of Vibrio cholerae and enterotoxigenic E.coli
XX CC simultaneously in a simple way.
XX CC The method has the same sensitivity as the dangerous radio-isotope
XX CC detection method.
XX S0 Sequence 1148 BP; 399 A; 205 C; 229 G; 315 T; 0 other:

```

```

alignment_scores:
    Quality: 208.00      Length: 208
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x AA023864 ..

```

```

Align seg 1/1 to: AA023864 from: 1 to: 1148

```

```

5 LeuTYrARgAlaAspSerArpProAspGluIleLysArGserGlygl 21
|||||
67 TTATACCGTGGCTGACTCTAGACCCGACATGAATAAAGCTTCCGAGG 116
|||||
21 yLeuMetProArGgLYHIsAnGluTYrPhAspArGgLYThGlMeTa 38
|||||
117 TCTTATGCCAGAGGCGCATATGAGTACTTCGATAGAGGAACCTCAATGA 166
|||||
38 snIleAsnLeuTYrAspHIsAlaArGgLYThGlnThrGlyPheValArg 54
|||||
167 ATATTATCTTTATGTATCATCCGACGAGAGAACACAAACCGGCTTTCACGA 216
|||||
55 TYrAspAspGlyTYrValSerThrSerLeuSerLeuArGserAlaHisLe 71
|||||
217 TATGATGACGAGATATGTTCCACTCTCTTACTTGAAGTCTCAGT 266
|||||
71 uAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYrTYrIleTYrValI 88
|||||
267 AGCAGAGCAGTCATATTATTCAGAGATATCCACTTACTATATATATGCTTA 316
|||||
88 leAlaThrAlaProAsnMetPheAsnValAsnAspValleuGlyValTYr 104
|||||
317 TAGCGACAGACCAATATATGTTTATATGATATGATATAGCGGATATAC 366
|||||
105 SerProHisProTYrGluGlnGluValSerAlaLeuGlyGlyIleProTY 121
|||||
367 AGCCCTACACCATATGATACAGAGAGGTTTCGCGTTAGTGGATATACATA 416
|||||
121 rSerGlnIleTYrGlyTYrTYrArGValAsnPhGlyValIleAspGluA 138
|||||
417 TTCTCAGATATATGATGATGATGCTGTATTTTGGTGTGATGATGATGAC 466

```

```

138 rPLeuHIsArGAsnArGgLUtTYrARgAspArGTyTYrArGAsnLeuAsn 154
|||||
467 GATTACATCTGTAACAGAGGAATATAGAACCGGTAATTACAGAAATCTGAAT 516
|||||
155 IleAlaProAlaGluAspGlyTYrARgLeuAlaGlyPheProProAspH 171
|||||
517 ATAGCTCCGGCAGAGAGATGTTACAGATTAGCAGAGTTTCCACCGGATCA 566
|||||
171 sGlnAlaTPArGgLUgluProTPHleHIsAlaProGlnGlyCysG 188
|||||
567 CCAAGCTTGAGAGAGAAACCTGATTCATCATGACCAAGAGTTGTG 616
|||||
188 lYAsnSerSerArGThrIleThrGlyAspPhrCysAnGluTYrThGln 204
|||||
617 GAAATTCATCAAGACATATACAGGTGATCTTGTATATGAGAGACCCAG 666
|||||
205 AsnLeuSerThrIleTYrLeuArg 212
|||||
667 AATCTGACGACATATATATCTCAGG 690
|||||
seq_name: /SID5/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51106
seq_documentation_block:
ID AAA51106 standard; cDNA; 782 BP.
XX
XX AAA51106;
XX
XX 26-SEP-2000 (first entry)
XX
XX DE Plant-optimized E. coli heat labile toxin A subunit coding sequence.
XX
XX KM Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
XX KM adjuvant; anti-bacterial; ss.
XX OS Escherichia coli.
XX
XX Key Location/Qualifiers
XX CDS 3..782
XX FT /*tag= a
XX FT /product= heat-labile-toxin-A
XX FT sig_peptide 3..56
XX FT /*tag= b
XX FT mat_peptide 57..779
XX FT /*tag= c
XX
XX WO200037609-A2.
XX
XX PD 29-JUN-2000.
XX
XX PF 22-DEC-1999; 99WO-US30747.
XX
XX PR 22-DEC-1998; 98US-0113507.
XX
XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
XX PA (MASO/) MASON H S.
XX PA (ARNT/) ARNTZEN C J.
XX
XX PI Mason HS, Arntzen CJ;
XX
XX DR WPI; 2000-442653/38.
XX DR P-PSDB; AA96646.
XX
XX PT New polynucleotides encoding LT-A or CT-A polypeptides for the
XX PT transformation of plant cells, useful in immunogenic compositions to
XX PT elicit immune responses in animals
XX
XX PS Example 1; Fig 1; 103pp; English.
XX
XX CC This plant-codon optimized cDNA encodes a synthetic Escherichia coli
XX CC heat-labile toxin (LT) A subunit (LT-A). The sequence contains
XX CC plant-preferred codons and eliminates sequence motifs associated with

```


CC spurious mRNA processing. A single codon insertion (GNG encoding valine)
 CC was made to accommodate the creation of a NcoI restriction site around the
 CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
 CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
 CC (CT-A) polypeptide, which have reduced enzyme activity as compared to the
 CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
 CC is altered to a plant preferred codon. The polynucleotide further
 CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
 CC subunit (CT-B). The polynucleotides are useful for the transformation of
 CC plant cells for the production of transgenic plants to produce edible
 CC vaccines, especially oral vaccines in transgenic plants for the
 CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
 CC mutant polypeptides are also useful as adjuvants.

CC Sequence 782 BP; 225 A; 186 C; 191 G; 180 T; 0 other;

alignment_scores:

Quality: 184.00 Length: 184
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA51106 ..

Align seg 1/1 to: AA51106 from: 1 to: 782

```

5 LeuTYrARgAlAsPSeArGPrOAsPGLuIlElySaRSeRgLygl 21
|||||
72 CTCTATrAGGGCTGACTGTAGGCCCCAGATGATCAAGAGCTGTGAGG 121
|||||
21 yLeuMeTPrOArGlyhIsAnGluTYrPhEAsPaRgLyThGlInMeTa 38
|||||
122 TCTCATGCGCCAAAGGGGACACAAATGACTTTGATAGGGGAACCAATGA 171
|||||
38 snlEAsnLeuTYrAsPhIsAlArGlyThRgnThRgLyPhEValArG 54
|||||
172 ACATCAACCTCTATGACCATGTAGGGGAACCAACTGGCTTTGTGAGG 221
|||||
55 TYrAsPaSPGLyTYrValSerThrSerLeuSerLeuArSeRaLhIsLe 71
|||||
222 TATGATGATGATATGTCTCACCCTCCTAGCTTGAAGCTCTGCTCACTT 271
|||||
71 uAlaGlyInSerlLeuSerGlyTYrSerThrTYrTYrIleTYrValI 88
|||||
272 GGCAGGACATFCATCCTCTCAGGATACCTCACCTACTACATCATGTGA 321
|||||
88 lEalArThrAlaPrOAsnMeTPhEAsnValAsnAspValIlEugLyValTYr 104
|||||
322 TTGCTTACAGCACCAACATGTCTCAATGTGAATGTGTGGGAGTGTAT 371
|||||
105 SerProHIsProTYrGInGInGluValSerAlaLeuGlyLyIleProTY 121
|||||
372 AGGCTTACCCATATGAGCAAGAGTCTGTGCTTGGGTGGAATCCATA 421
|||||
121 rSeRgInlIeTYrGlyTYrTYrArGValAsnPhEgLyValIlEAspGLuA 138
|||||
422 CTCCCAATCTATGATGTATAGGGGGAACCTTGGGTGATGATGAGA 471
|||||
138 rglEuhIsArGAsnArGluTYrArGAsPaRgTYrTYrArGAsnLeuAsn 154
|||||
472 GGGCTCCATAGGAATAGGAGTATAGGACAGGTACTATAGAACCTTAC 521
|||||
155 lAlaAlProAlaGluAspGLyTYrArGLeuAlaGlyPhEProProAspHl 171
|||||
522 ATAGCTCCACAGAGAGATGTTATAGTTGGCAGGTTTCCACACAGACCA 571
|||||
171 sGlnAlATrPaRgGluInuProTYrIleHIsAlaProGInGlyCysG 188
|||||
572 CCAAGCTGTGAGGAGAGCCCTGTGATCCACATGCACCAAGGTTGTG 621
|||||
188 Ty 188
||

```

622 GA 623

seq_name: /SID85/gcdata/geneseq/geneseqn_emb1/NA2000.DAT:AA51544

seq_documentation_block:

ID AA51544 standard; cDNA: 782 BP.

XX AA51544;

DT 26-SEP-2000 (first entry)

DE Plant-optimized E. coli LT-A G192 mutant coding sequence.

KW Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;

KW adjuvant; anti-bacterial; R192G; ss.

OS Escherichia coli.

OS Synthetic.

Key Location/Qualifiers

FT CDS

FT sig_peptide

FT mat_peptide

FT

WO200037609-A2.

PD 29-JUN-2000.

PF 22-DEC-1999;

PR 22-DEC-1998;

PA (BOYC-) BOYCE THOMPSON INST PLANT RES.

PA (MASO-) MASON H S.

PA (ARNT/) ARNTZEN C J.

PI Mason HS, Arntzen CJ;

DR WPI: 2000-442653/38.

DR P-PSDB: AAY96648.

PT New polynucleotides encoding LT-A or CT-A polypeptides for the

PT transformation of plant cells, useful in immunogenic compositions to

PT elicit immune responses in animals

XX Example 3; Page -: 103pp; English.

XX This cDNA encodes a mutant R192G Escherichia coli heat-labile toxin (LT)

CC A subunit (LT-A). The nucleotide at position 631 was changed from the

CC wild-type A to G which causes a change of arginine to glycine in the

CC mature protein at residue 192. The sequence contains plant-preferred

CC codons and eliminates sequence motifs associated with spurious mRNA

CC processing. A single codon insertion (GNG encoding valine) was made to

CC accommodate the creation of a NcoI restriction site around the

CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A

CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit

CC (CT-A) polypeptide, which have reduced enzyme activity as compared to

CC the wild-type LT-A or CT-A polypeptide and where at least one of the

CC codons is altered to a plant preferred codon. The polynucleotide further

CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B

CC subunit (CT-B). The polynucleotides are useful for the transformation of

CC plant cells for the production of transgenic plants to produce edible

CC vaccines, especially oral vaccines in transgenic plants for the

CC prophylactic or therapeutic treatment against E. coli or V. cholerae.

CC Note: This sequence does not appear in the specification. It was

CC constructed from the wild type LT-A cDNA shown in AA51106 which is

CC given in Figure 1 of the specification.

sq Sequence 782 BP; 224 A; 186 C; 192 G; 180 T; 0 other;

alignment_scores:

Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAN51544 ..

Align seg 1/1 to: AAN51544 from: 1 to: 782

```

5 LeuTYrArgAlaAspSerArgProProAspGluLeuLysArgSerGly 21
72 CTCTATAGGCTGACTCTAGAGCCCAAGATGATGACAAAGAGTCTGAGG 121
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
122 TCTCATGCCAAGGGGACACATAGTACTTTGATAGGGGAACCAATGA 171
38 snlLeaSnLeuTYrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
172 ACATCAACCTCTATGACCATGCTAGGGGAACCCAACTGGCTTTGTGAGG 221
55 TyrAspAspGlyTYrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
222 TATGATATGATATGATATGTCACCTCCCTTATGATGAGTGTCTCACT 271
71 uAlaGlyLysSerIleLeuSerGlyTYrSerThrTYrTYrIleTYrValI 88
272 GGCGAGGACATTCATCCTCTCAGGATCTCCACCTACTACTATCATGTA 321
88 leaIaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTYr 104
322 TTGGCTACAGACCAACATGTTCAATGTGATGTGTGGGAGTGTAT 371
105 SerProHisProTYrGluGlnGluValSerAlaLeuGlyGlyIleProTY 121
372 AGCCCTACCCATATGAGCAAGAGGTCTCTGCTTGGGTGGAAATCCCAT 421
121 rSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyValIleAspGlu 138
422 CTCCTCAATCATATGATGATGATAGGTGTAAGTCTTGTGTATGATGA 471
138 rGleuHisArgAsnArgGluTYrArgAspArgTYrTYrArgAsnLeuAsn 154
472 GGCTCCATAGGAATAGGAGATATAGGACAGGTACTATAGCAACCTCAAC 521
155 lLeaIaProAlaGluAspGlyTYrArgLeuAlaGlyPheProProAspHI 171
522 ATAGCTCCAGCAGAGAGATGCTTATAGGTGGCAGGTTTCCACCAAGACA 571
171 sGlnAlaTYrParGluGluProTYrIleHisHisAlaProGlnIleCysG 188
572 CCAAGCCTGGAGGAGAGGCCCTGATCCACCATGACACACAGAGTTGTG 621
188 y 188
622 GA 623
seq_name: /STDS/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: AAN50206
seq_documentation_block:
ID AAN50206 standard; DNA; 777 BP.
XX
AC AAN50206;
XX
DT 30-OCT-1991 (first entry)
XX
DE Sequence of the pig scours heat labile toxin (LT) LTA gene of the
XX site directed mutant SDM1.
XX
```

KW Pig scours vaccine; toxin; diarrhoea; ss.

XX E.coli NCIB 11932.

OS Key Location/Qualifiers

FT CDS 1..777

FT misc_difference 235..237

FT /tag= b

FT /note= "TCC in native sq"

XX EP145486-A.

XX 19-JUN-1985.

XX 12-DEC-1984; 84EP-0308620.

XX 12-DEC-1983; 83GB-0033131.

XX (GLAX) GLAXO GROUP LTD.

XX Hayes MV, Harford S, Ross GW;

XX WPI; 1985-148358/25.

XX P-PsDB; AAP50191.

XX New toxoid as inactivated form of toxin for use in vaccines - 1s

XX obd. from organism transformed by gene

XX Example; Fig 2; 61pp; English.

XX AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of the site directed mutant SDM1 (see AAN50206) is inactive. The inventors claim a vaccine prepn. active against pig scours which contains an inactivated LTA component, together with additional K88 antigens opt. with whole cells comprising the antigens or contg. the inactivated LTA.

XX Sequence 777 BP; 255 A; 135 C; 164 G; 223 T; 0 other;

alignment_scores:

Quality: 179.00 Length: 179
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAN50206 ..

Align seg 1/1 to: AAN50206 from: 1 to: 777

```

62 ThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
238 ACTTCTCTTACTTGTAGAGAGTGCACACTTACAGACAGTCTAATATATC 287
78 rGlyTYrSerThrTYrTYrIleTYrValIleAlaThrAlaProAsnMet 95
288 AGGATATCCACTTACTATATATATATATAGCGACAGACCAATATATGT 337
95 heaSnValAsnAspValLeuGlyValTYrSerProHisProTYrGluGln 111
338 TTATATGTTAATGATGATATAGCGGATATACAGCCSCACCCATATGAGAG 387
112 GluValSerAlaLeuGlyGlyIleProTYrSerGlnIleTYrGlyTYrPT 128
388 GAGCTTCTCGCTTAGGAGATACCAATATCTCAGATATGTGATGGTA 437
128 rArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluT 145
438 TCGTGTATATTTGTGTGATGATGAGAACGATTCATCGTATACAGGAGAT 487
145 yArgAspArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspGly 161
|||||
```


seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA501506
 seq_documentation_block:
 ID AA501506 standard; DNA; 1508 BP.
 XX
 AC AA501506;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE E. coli heat-labile enterotoxin (LT) mutant LTdel110/112 DNA.
 XX
 KM Heat-labile enterotoxin; LT; LT63Y; LTdel110/112; mutant;
 KM detoxified and immunologically active protein; ADP-ribosylation; Gs;
 KM endotoxin; diarrhoea; ds.
 XX
 OS Escherichia coli strain K88ac.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FH CDS 164..1305
 FT /*tag= a
 FT /product= "LTdel110/112 mutant protein"
 FT /transl_except= (pos:932..933,aa:Met)
 FT /note= "This codon has an apparent 1 nucleotide
 FT deletion which alters the reading frame"
 FT
 XX
 PN WO200119998-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 15-SEP-1999; 99WO-KR00555.
 XX
 PR 15-SEP-1999; 99WO-KR00555.
 XX
 PA (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
 PI
 PI Park EJ, Kim JS, Chang J, Yum J, Chung S;
 XX
 DR WPI: 2001-281524/29.
 DR P-PSDB: AAU00507.
 XX
 XX New detoxified mutants of Escherichia coli heat-labile enterotoxin
 PT useful as vaccine for preventing and treating diarrhoea, and as adjuvant
 PT for antibody production -
 PT
 PS Claim 7; Page 45-46; 48pp; English.
 XX
 CC The present sequence encodes for Escherichia coli heat-labile
 CC enterotoxin (LT) mutant LTdel110/112. LT63Y (AAU00506) and LTdel110/112
 CC are two novel detoxified and immunologically active proteins (LT
 CC mutants) derived by site-directed mutagenesis of the A1 subunit of wild
 CC type LT. The substitution of Ser to Tyr at position 63 in LT63Y blocks
 CC NAD-binding. Deletion of Glu residues at positions 110 and 112 in
 CC LTdel110/112 eliminate the enzymatic activity of LT. The A1 subunit of
 CC wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
 CC regulates CAMP levels. The resulting increase in CAMP is the cause of
 CC diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
 CC of mutant heat-labile endotoxins LT63Y and LTdel110/112 were tested.
 CC Groups of mice were immunised with LT63Y or LTdel110/112. The control
 CC groups received phosphate buffered saline (PBS) alone. The serum and
 CC faecal antibody titres to LT were determined. The results showed that
 CC mice immunised with LT63Y or LTdel110/112 contained high and
 CC comparable level of anti-LT antibodies in sera and faecal extracts
 CC compared with those immunised with wild-type LT. The LT mutants are
 CC useful as a vaccine for preventing and treating diarrhoea and as an
 CC adjuvant for antibody production.
 XX
 SO Sequence 1508 BP; 506 A; 277 C; 295 G; 430 T; 0 other;

alignment_scores:
 Quality: 128.00 Length: 128
 Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x AA501506 ..
 Align seq 1/1 to: AA501506 from: 1 to: 1508
 113 ValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrTyr 129
 548 GTTCTCGGTAGGATGCAATATTCATGATATATGATGATGATCG 597
 129 GValAspPheGlyValIleAspGluArgLeuNIleArgAsnArgGlyTyr 146
 598 TGTATATTTTGTGTGATGATGAACGATTTCATCGTACAGGGAATTA 647
 146 rGaAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
 648 GAGACCGGTATTACAGAAATCTGAATATAGCTCCGCGAGAGATGCTTAC 697
 163 ArgLeuAlaGlyPheProProAspPheIleGlnAlaTyrArgGluIleProTr 179
 698 AGATTAGCAGGTTTCCACCGGATCCACCACTTGAGAGAAACCTG 747
 179 PIIleHisIleAlaProGlnGlyCysGlyAsnSerSerArgThrIlePhe 196
 748 GATTCATCATGCACACCAAGCTGTGGAAATTCATCAAGAACATCATCAG 797
 196 lYAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArg 212
 798 GTGATACTGTGTAAAGAGAGACCAAGATCTGACACAAATATATCTCAG 847
 213 GluTyrGlnSerIysValIysArgGlnIlePheSerIleTyrGlnSerG 229
 848 GAATATCATCAAAAGTTAAAGAGCAGATATTTTCAGCTATCAGTCAGA 897
 229 uValAspIleTyrAsnArgIleArgAspGluLeu 240
 898 GGTGACATATATATAACAGAAATTCGGGATGAATTA 931
 seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51147
 seq_documentation_block:
 ID AAA51147 standard; cDNA; 782 BP.
 XX
 AC AAA51147;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Plant-optimized E. coli LT-A K63 mutant coding sequence.
 XX
 KM Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
 KM adjuvant; anti-bacterial; S63K; ss.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FH CDS 3..782
 FT /*tag= a
 FT /product= heat-labile-toxin-A
 FT sig-peptide 3..59
 FT mat-peptide 60..779
 FT /*tag= c
 XX
 PN WO200037609-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 22-DEC-1999; 99WO-US30747.
 XX
 PR 22-DEC-1998; 98US-0113507.
 XX


```

164 LeuAIIgLYpHeProRoAsPhISglNaIATrPaRGInGUlPrOTfPl 180
|||||
549 TTGGAGGTTTCCACCAcCACCACAAGCTTGAGGAGAGGCCCTCGAT 598
|||||
180 eHISHIsAlAProGInGLyCySgLy 188
|||||
599 CCACCATGCACCAACAAGGTGTGA 623

seq_name: /SDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA51545
seq_documentation_block:
ID AAA51545 standard; cDNA; 782 BP.
XX
XX
AAAS1545;
AC
XX
DT 26-SEP-2000 (first entry)
XX
DE Plant-optimized E. coli LT-A R72 mutant coding sequence.
XX
KM Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
XX adjutant; anti-bacterial; A72R; ss.
XX
OS Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers
FH CDS 3..782
FT /*tag= a
FT /*product= heat-labile_toxin-A
FT sig_peptide 3..59
FT mat_peptide 60..779
FT /*tag= c
XX
PN W0200037609-A2.
PD
PD 29-JUN-2000.
XX
PF 22-DEC-1999; 99WO-US30747.
XX
PR 22-DEC-1998; 98US-O113507.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
XX Mason HS, Arntzen CJ;
XX
DR WP1: 2000-442653/38.
DR P-PSDB; AA956649.
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX Example 4; Page -; 103pp; English.
XX
XX This CDNA encodes a mutant A72R Escherichia coli heat-labile toxin (LT)
CC A subunit (LT-A). The nucleotides at position 273-274 were changed from
CC the wild-type GC to AG which causes a change of alanine to arginine in
CC the mature protein at residue 72. The sequence contains plant-preferred
CC codons and eliminates sequence motifs associated with spurious mRNA
CC processing. A single codon insertion (GTG encoding valine) was made to
CC accommodate the creation of a NcoI restriction site around the
CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
CC (CT-A) polypeptide, which have reduced enzyme activity as compared to
CC the wild-type LT-A or CT-A polypeptide and where at least one of the
CC codons is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible

```


CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae.
CC The mutant polypeptides are also useful as adjuvants.
CC Note: This sequence does not appear in the specification. It was
CC constructed from the wild type LT-A cDNA shown in AA051106 which is
CC given in Figure 1 of the specification.

XX Sequence 782 BP; 226 A; 185 C; 191 G; 180 T; 0 other;

alignment_scores:

Quality: 116.00 Length: 116
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051545 ..

Align seg 1/1 to: AA051545 from: 1 to: 782

```
73 GlnGlnSerIleLeuSerGlyTyrSerThrTyrTrpIleValIleAl 89
|||||
276 GGACAAATCCATCTCTCAGATCTCCACCTACTACATCTATGTGATTGC 325
89 aThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerP 106
|||||
326 TACAGCACCAACAATGTTTCATGTGATGATGTGGAGATGTTATACC 375
106 roHISpPOTyrgLInGInLValSerAlaLeuGlyGlyIlePOTyriser 122
|||||
376 CTCACCCCTATGACCAAGAGCTGTCTGGTGGGGAATCCCACTACCTCC 425
123 GlnIleTyrgLInPOTyrgValAsnPheGlyValIleAspGluArle 139
|||||
426 CAATCTATGATGATGATAGGGTGAACCTTGTGTGATGATGAGAGCT 475
139 uHISArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleA 156
|||||
476 CCATAGAGATAGGAGATAGGAGAGAGTACTATAGGAACCTCAACATAG 525
156 IAProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGln 172
|||||
526 CTCACGACGAGAGATGTTATAGTGGCAGGTTTCCACACACACCAAA 575
173 AlArPrArGluGluProTrpIleHisIleAlaProGlnGlyCysGly 188
|||||
576 GCCTGGAGGAGAGAGCCCTGGATCCACCATGACACCAAGGTTGTGCA 623
```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051314

seq_documentation_block:

ID AA051314 standard; DNM: 711 BP.

XX AA051314;

XX 08-DEC-1993 (first entry)

XX Encodes Asp-53 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KM ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX OS

XX key Location/Qualifiers

FT misc_difference 153..155

FT /tag a

XX /note= "wild-type GTC(Val) mutated to GAT(Asp)"

XX MO09313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-0MI3513.

XX (BIOC-) BIOGINE SCLAVO SPA.

XX Domenighini M, Hol W, Piazza M, Rappuoli R;

XX WPI; 1993-227320/28.

XX P-PDB; AAR38730.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX - useful as vaccines against infection by Vibrio cholerae and

XX enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051314 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 235 A; 122 C; 156 G; 198 T; 0 other;

alignment_scores:

Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051314 ..

Align seg 1/1 to: AA051314 from: 1 to: 711

```
103 ValTyrSerProHisPOTyrgLInGInLValSerAlaLeuGlyIle 119
|||||
295 GTATCAGACCCCTCACCATATGACAGAGGTTTCTGCTTAGGTGGAAT 344
119 ePOTyriserGlnIleTyrgLInPOTyrgValAsnPheGlyValIleA 136
|||||
345 ACCATATCTCTCGATATATGATGATGATGATGATGATGATGATGATG 394
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGAACGATTACATGCTAACAGGGAATATAGACCGGTTATACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||
445 CTGAATATAGCTCCGCGCAGAGGATGTTACAGATTATAGCAGGTTTCCAC 494
169 oAspHisGlnAlaTrpArgGluGluProTrpIleHisIleAlaProGln 186
|||||
495 GGATCACCAGAGTTGGAGAGAAAGAACCTGATTCATCATGACCAACAAG 544
186 IyCysGly 188
|||||
545 GTTGTGGA 552
```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051315

seq_documentation_block:

ID AA051315 standard; DNM: 711 BP.

XX

XX AA051315;


```

DT 08-DEC-1993 (first entry)
XX
DE Encodes Glu-53 E.coli heat labile toxin subunit A.
XX
XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KM protome A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT misc_difference 153..155
FT /tag=^a
FT /note= "wild-type GTC(Val) mutated to GAA(Glu)"
XX
PN W09313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91TT-OMI3513.
XX
PA (BIOC-) BIOCINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Pizza M, Rappuoli R;
XX
DR WPI: 1993-227320/28.
DR P-PSDB; AAR38731.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
XX
PS Claim 3; Fig 2 and Page 46; 60pp; English.
XX
CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AAQ51314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AAQ51315 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 BP; 236 A; 122 C; 156 G; 197 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ51315 ..
Align seg 1/1 to: AAQ51315 from: 1 to: 711

103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIle 119
|||||
295 GTATACAGCCCTCACCCATATGAACAGAGGTTCTCGTTAGCGTGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAlaPheGlyValIleA 136
|||||
345 ACCGATATTCAGATATATGATGATCGTTAATTTTGGTGTGATG 394
|||||
136 SPGIUAAGLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACGATTAACATCGTACACAGGAATATAGAGCCGTTATACAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169

```

```

|||||
445 CTGAAATATACCTCCGACAGAGATGTTACAGATTACAGGTTCCACC 494
|||||
169 oAspHisGlnAlaTyrPargGluGlnProTyrPheHisAlaProGln 186
|||||
495 GGATCCACCAAGCTTGGAGAGAAAGACCTGGATTCATCATGCACACAG 544
|||||
186 LysGly 188
|||||
545 GTTGTGCA 552

seq_name: /STD5/gcgcdata/geneseq/geneseqn_emb1/NA1993.DAT:AAQ51316
seq_documentation_block:
ID AAQ51316 standard; DNA; 711 BP.
XX
XX AAQ51316;
AC
XX
XX
DT 08-DEC-1993 (first entry)
XX
DE Encodes Tyr-53 E.coli heat labile toxin subunit A.
XX
XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KM protome A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT misc_difference 153..155
FT /tag=^a
FT /note= "wild-type GTC(Val) mutated to TAC(Tyr)"
XX
PN W09313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91TT-OMI3513.
XX
PA (BIOC-) BIOCINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Pizza M, Rappuoli R;
XX
DR WPI: 1993-227320/28.
DR P-PSDB; AAR38732.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
XX
PS Claim 3; Fig 2 and Page 46; 60pp; English.
XX
CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AAQ51314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AAQ51316 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 BP; 235 A; 123 C; 155 G; 198 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x AAQ51316 ..

Align seg 1/1 to: AAQ51316 from: 1 to: 711

```
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyI119
|||||
295 GTATACAGCCCTCACCATATGAAACAGAGGTTCTCGCTTAGTGGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyVal111ea 136
|||||
345 ACCATATTCTCAGATATATGATGATCGTATCGTAAATTTGGTGTGATTCG 394
|||||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACACATTATCATCGTACACGGGAATATGAGACCGGTATTCAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProP 169
|||||
445 CTGAATATAGCTCCGGCAGAGATGTTACAGATTACAGAGTTCCACACC 494
|||||
169 OASPHISGlnAlaTrrParGluGluProTrrPLeHisHisAlaProGln 186
|||||
495 GGATCACCAAGCTTGAGAGAGAAGAACCTGGATTCATCATGCACACAA 544
|||||
186 LYGSGLY 188
|||||
545 GTTGTGGA 552
```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51317

seq_documentation_block:

ID AAQ51317 standard; DNA; 711 BP.

AC AAQ51317;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-63 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers

misc_difference 184..186

/*tag= a

/note= "wild-type TCT(Ser) mutated to AAG(Lys)"

WO9313202-A.

08-JUL-1993.

30-DEC-1992; 92WO-EP03016.

31-DEC-1991; 91TT-OMI3513.

(BIOC-) BIOCINE SCLAVO SPA.

Domenighini M, Hol W, Pizzera M, Rappuoli R;

WPI: 1993-227320/28.

P-PSDB; AAR44016.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AAQ51314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AAQ51317 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

alignment_scores:

Quality:	86.00	Length:	86
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AAQ51317 ..

Align seg 1/1 to: AAQ51317 from: 1 to: 711

```
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyI119
|||||
295 GTATACAGCCCTCACCATATGAAACAGAGGTTCTCGCTTAGTGGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyVal111ea 136
|||||
345 ACCATATTCTCAGATATATGATGATCGTATCGTAAATTTGGTGTGATTCG 394
|||||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACACATTATCATCGTACACGGGAATATGAGACCGGTATTCAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProP 169
|||||
445 CTGAATATAGCTCCGGCAGAGATGTTACAGATTACAGAGTTCCACACC 494
|||||
169 OASPHISGlnAlaTrrParGluGluProTrrPLeHisHisAlaProGln 186
|||||
495 GGATCACCAAGCTTGAGAGAGAAGAACCTGGATTCATCATGCACACAA 544
|||||
186 LYGSGLY 188
|||||
545 GTTGTGGA 552
```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51318

seq_documentation_block:

ID AAQ51318 standard; DNA; 711 BP.

AC AAQ51318;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-97 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers

misc_difference 277..279

/*tag= a

/note= "wild-type GTT(Val) mutated to AAG(Lys)"

WO9313202-A.

08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.
 XX 31-DEC-1991; 91IT-0MI3513.
 XX (BIOC-) BIOCINE SCLAVO SPA.
 PA Domenighini M, Hol W, Piazza M, Rappuoli R;
 PI WPI; 1993-227320/28.
 DR P-PSDB; AAR44017.
 XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051318 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 236 A; 123 C; 156 G; 196 T; 0 other;

alignment_scores:
 Quality: 86.00 Length: 86
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051318 ..

Align seg 1/1 to: AA051318 from: 1 to: 711

```

103 VAllyrSerPrOHIsPrOTyrglIngluValSerAlaLeuGlyYlI 119
|||||
295 GTATACAGCCCTCACCCTATATGACAGAGAGGTTCTCGGTAGGTGAT 344
|||||
119 ePrOTySerGlnIleTyrglYTrPTyrgValAsnPhelGlyValIleA 136
|||||
345 ACCATATTCTCAGATATATGATGATGATCGTGAATTTTGGTGATG 394
|||||
136 sPgluArGleuHIsArGAsnArGluTyrgAspArGTYrTYrArGAsn 152
|||||
395 ATGAACGATTACATCGTAACAGGAAATAGAGACCGGATATACGAAT 444
|||||
153 LeuAsnIleAlaPrOAlaGluAspGlyTyrgValAlaGlyPhePrOT 169
|||||
445 CTGAATATATAGCTCCGCGAGAGATGTTACAGATTAGCAGAGTTTCCAC 494
|||||
169 cAspHIsGlnAlaTrPARgGluGluPrOTrIleHIsAlaPrOGing 186
|||||
495 GGATACACCAAGCTTGAGAGAGAAACCTGGATTATCATCACCACAAAG 544
|||||
186 lYcYsGly 188
|||||
545 GTTGTGA 552

```

seq_name: /SIDS5/9cgcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051319

seq_documentation_block:

ID AA051319 standard; DNA; 711 BP.
 XX
 AC AA051319;
 XX
 DT 08-DEC-1993 (first entry)

XX
 DE Encodes Tyr-97 E.coli heat labile toxin subunit A.
 XX
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
 KW protomer A; site-directed mutagenesis; reduced toxicity;
 KW ADP-ribosyltransferase activity; ss.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 277..279
 FT /tag= a
 FT /note= "wild-type GGT(Val) mutated to TAC(Tyr)"
 XX
 PN MO9313202-A.
 XX
 PD 08-JUL-1993.
 XX
 XX 30-DEC-1992; 92WO-EP03016.
 PF 31-DEC-1991; 91IT-0MI3513.
 XX
 PR (BIOC-) BIOCINE SCLAVO SPA.
 XX
 PA Domenighini M, Hol W, Piazza M, Rappuoli R;
 XX
 PI WPI; 1993-227320/28.
 DR P-PSDB; AAR44018.
 DR Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051319 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 235 A; 124 C; 155 G; 197 T; 0 other;

alignment_scores:
 Quality: 86.00 Length: 86
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051319 ..

Align seg 1/1 to: AA051319 from: 1 to: 711

```

103 VAllyrSerPrOHIsPrOTyrglIngluValSerAlaLeuGlyYlI 119
|||||
295 GTATACAGCCCTCACCCTATATGACAGAGAGGTTCTCGGTAGGTGAT 344
|||||
119 ePrOTySerGlnIleTyrglYTrPTyrgValAsnPhelGlyValIleA 136
|||||
345 ACCATATTCTCAGATATATGATGATGATCGTGAATTTTGGTGATG 394
|||||
136 sPgluArGleuHIsArGAsnArGluTyrgAspArGTYrTYrArGAsn 152
|||||
395 ATGAACGATTACATCGTAACAGGAAATAGAGACCGGATATACGAAT 444
|||||
153 LeuAsnIleAlaPrOAlaGluAspGlyTyrgValAlaGlyPhePrOT 169
|||||

```



```

445 CTGATATAGCTCGGCGAGAGATGGTTACAGATTAGCAGGTTTCCACC 494
169 OASPHISGLNALATRPARGIUGLUPROTPRIEHSIALAproGing 186
|||||
495 GGATCACCACCTTGAGAGAAACCTGGATTATCATGCACACCAAG 544
186 IYCysGly 188
|||||
545 GTTGTGGA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ42768

seq_documentation_block:
ID AAQ42768 standard; DNA; 711 BP.
XX
AC AAQ42768;
XX
DT 08-DEC-1993 (first entry)
XX
DE E.coli heat labile toxin subunit A coding sequence.
XX
KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KM ADP-ribosyltransferase activity; ss.
XX
OS Escherichia coli.
FH
FT key Location/Qualifiers
FT mat_peptide 1..708
FT /*tag= a
FT /note= "LT-A"
XX
PN WO9313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91IT-OMI3513.
XX
PA (BIOC-) BIOGINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Piazza M, Rappuoli R;
XX
DR WPI: 1993-227320/28.
DR P-PSDB; AAR38728.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
PS Disclosure; Fig 2; 60pp; English.
XX
CC This sequence encodes the A subunit of the heat labile toxin (LT-A)
CC of a strain of E.coli known to affect humans. The sequence was
CC published by Yamamoto et al. J Biol. Chem. 259, 5037-5044.
CC Mutations at selected positions within this sequence have been found
CC to reduce toxicity (see AAQ51314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli.
XX
SQ Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ42768 ..
Align seg 1/1 to: AAQ42768 from: 1 to: 711

```

```

103 VALTYSerPRONISPROTYrGLUGInGLuValSerAlaLeuGLyYL 119
|||||
295 GTATACACCCCTCACCCATATGACAGAGGTTTCTGCGTTAGGTGAAT 344
119 ePROTYSerGlnILETYrGLYrTPYrArgValAsnPhGlyVal11lea 136
|||||
345 ACCATATTCTCAGATATATGAGTGTATCGTGTATTTGGTGTGATTG 394
136 SPGLUArgLeuNHISArgAsnArgGLUTYrArgSPArgTYrTYrArgAsn 152
|||||
395 ATGAACGATTTACATGCTAAACAGGAAATATAGACCGGTATTACAGAAAT 444
153 LeuAsnILEAlaPRoAlaGLuAspGLYrTYrArgLeuAlaGLYrPhProPr 169
|||||
445 CTGATATAGCTCCGCGAGAGATGGTTACAGATTAGCAGGTTTCCACC 494
169 OASPHISGLNALATRPARGIUGLUPROTPRIEHSIALAproGing 186
|||||
495 GGATCACCACCTTGAGAGAAACCTGGATTATCATGCACACCAAG 544
186 IYCysGly 188
|||||
545 GTTGTGGA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV81595

seq_documentation_block:
ID AAV81595 standard; DNA; 711 BP.
XX
AC AAV81595;
XX
DT 11-MAY-1999 (first entry)
XX
DE E. coli heat labile toxin gene.
XX
KW A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;
KW parenteral adjuvant; antigen; immunisation; humoral response;
KW cell-mediated immune response; virus; bacterium; parasite; fungus;
KW tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;
KW systemic lupus erythematosus; Alzheimer's disease; diagnosis; ss.
XX
OS Escherichia coli.
FH
FT key Location/Qualifiers
FT mat_peptide 1..708
FT /*tag= a
FT /note= "LT-A"
XX
PN WO9842375-A1.
XX
PD 01-OCT-1998.
XX
PF 19-MAR-1998; 98MO-US05454.
XX
PR 18-MAR-1998; 98US-0044696.
PR 21-MAR-1997; 97US-0041227.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Barchfeld G, Del Giudice G, Rappuoli R;
XX
DR WPI: 1999-070064/06.
DR P-PSDB; AAM67772.
XX
PT Detoxified mutants of bacterial ADP-ribosylating toxins as
PT parenteral adjuvants - useful to enhance humoral and cell-mediated
PT immune responses in vertebrates when administered with selected
PT antigen e.g. in disease treatment
XX
PS Disclosure; Fig 1A-B; 51pp; English.
XX
CC This sequence corresponds to the coding region for the A subunit of the
CC E. coli heat labile toxin, an example of a bacterial ADP-ribosylating
CC toxin. A mutant detoxified form of this protein is used in a parenteral
CC adjuvant composition, which comprises the detoxified protein, at least
CC one selected antigen and optionally a pharmaceutically acceptable
CC (optionally topical) vehicle. The adjuvant composition can be

```


CC administered parenterally in conjunction with at least one antigen in
CC methods to immunise vertebrate subjects. The adjuvant has the ability
CC to enhance the humoral and cell-mediated immune responses elicited by
CC the antigen (e.g. by making the antigen more strongly immunogenic or
CC necessitating fewer/lower antigen doses). It can be administered
CC prior/subsequent to the antigen, and is preferably administered within
CC a short space of time to the same site; it can also be administered in
CC isolation from antigens as a boost following systemic or mucosal antigen
CC administration. Most preferably, the adjuvant is co-administered with
CC the antigen in the compositions and a pharmaceutically acceptable
CC carrier. The antigen may be derived from viruses, bacteria, parasites
CC and fungi or may be tumour antigens, self-antigens and allergens. The
CC compositions are therefore useful in the treatment and prevention of
CC e.g. viral diseases, allergic manifestations, diseases caused by
CC pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases
CC (e.g. Systemic Lupus Erythematosus), Alzheimer's disease and cancers.
CC The adjuvant can also be used to prepare antibodies against selected
CC antigen(s), useful e.g. for diagnostic purposes or for antigen
CC purification.

Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAV81595 ..

Align seg 1/1 to: AAV81595 from: 1 to: 711

```
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIle 119
|||||
295 GATACAGCCCTCCACCAATGACAGAGAGTTTCGCGTAGGTGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrTyrValAsnPheGlyValIle 136
|||||
345 ACCATATTCACAGATATATGATGATCGTATCGTTAATTTGGTGATG 394
|||||
136 SPGLuAlaGLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArg 152
|||||
395 ATGACACATTACATCGTAAACAGGAAATATAGAGCCGTTATTCAGAA 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgPheAlaGlyPhePro 169
|||||
445 CTGAATATACCTCCGCGACAGAGATGTTACAGATTACAGGTTCCACC 494
|||||
169 OASPHISGlnAlaTPrArgGluGluProTPrIleHisAlaProGln 186
|||||
495 GGATCACCAAGCTTGAGAGAGAACCCGTGATTCATCATGCACCAAG 544
|||||
186 LysGly 188
|||||
545 GTTGTGGA 552
```

seq_name: /sids5/gcsgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051321

seq_documentation_block:

ID AA051321 standard; DNA; 711 BP.

XX AA051321;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers
FH misc_difference 298..300
FT /tag= a
FT /note= "wild-type TrAc(Tyr) mutated to AAc(Lys)"

XX W09313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-OM13513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizza M, Rappuoli R;

XX WPI; 1993-227320/28.

XX P-PSDB; AAR44020.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol.Chem., 259, 5037-5044 - see AA042766) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051321 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (amino
CC acid numbering is based on the cholera toxin A subunit sequence).

Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other;

alignment_scores:
Quality: 84.00 Length: 84
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051321 ..

Align seg 1/1 to: AA051321 from: 1 to: 711

```
105 SerProHisProTyrGluGluValSerAlaLeuGlyIleProTy 121
|||||
301 AGCCCTCACCAATATGACAGAGAGTTTCGCGTAGGTGAATACATA 350
|||||
121 rSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAsp 138
|||||
351 TTTCACATATATGATGATGATCGTAAATTTGGTGATGATGATGAC 400
|||||
138 rGLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 154
|||||
401 GATTACATCGTAAACAGGAAATATAGAGCCGTTATTCAGAAATCTG 450
|||||
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProAsp 171
|||||
451 ATGACCTCCGCGACAGAGATGTTACAGATTAGCAGATTCCACCGGAT 500
|||||
171 sGlnAlaTPrArgGluGluProTPrIleHisAlaProGlnLysG 188
|||||
501 CCAAGCTTGAGAGAGAACCCGTGATTCATCATGCACACCAAGTGTG 550
|||||
188 Lys 188
||
```


551 GA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051322

seq_documentation_block:

ID AA051322 standard; DNA; 711 BP.

AC AA051322;

DT 08-DEC-1993 (first entry)

DE Encodes Asp-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KM protomer A; site-directed mutagenesis; reduced toxicity;

OS Escherichia coli.

FH Key location/Qualifiers

FT misc_difference 298..300

FT /*tag= a /note= "wild-type TAC(Tyr) mutated to GAT(Asp)"

PD WO9313202-A.

PN 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Piazza M, Rappuoli R;

DR MPI: 1993-227320/28.

DR P-PSDB; AAR44021.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PS enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051322 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

SQ Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

alignment_scores:

Quality: 84.00

Ratio: 1.000

Length: 84

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051322 ..

Align seg 1/1 to: AA051322 from: 1 to: 711

105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyGlyProTyr 121
 111
 301 AGCCCTCCACCATATGAAACAGAGAGGTTTCGCGTTAGTGGAATACATA 350

121 rSerGlnIleTyrGlyTyrTyrArgValAsnPhnGlyValIleAspGlu 138
 138 rGluHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
 351 TTCTCAGATATATGAGATGATCGTCTTATTGTCGTGATGATGATAC 400
 401 GATTACATCGTAAACAGGAGATATAGACCGGATATACAGAAATCTGAAT 450
 155 ILeAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPheAsp 171
 451 ATAGCTCCGGCAGAGATGGTTACAGATTAGCAGGTTCCACCGGATCA 500
 171 sGlnAlaTyrArgGluGlnProTyrPheHisAlaProGlnGlyCysG 188
 501 CCAAGCTTGAGAGAGAACACCGTGGATTCATCATGACCAAGGTTGTG 550
 188 Iy 188
 551 GA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051323

seq_documentation_block:

ID AA051323 standard; DNA; 711 BP.

AC AA051323;

DT 08-DEC-1993 (first entry)

DE Encodes Ser-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KM protomer A; site-directed mutagenesis; reduced toxicity;

OS Escherichia coli.

FH Key location/Qualifiers

FT misc_difference 298..300

FT /*tag= a /note= "wild-type TAC(Tyr) mutated to AGT(Ser)"

PD WO9313202-A.

PN 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Piazza M, Rappuoli R;

DR MPI: 1993-227320/28.

DR P-PSDB; AAR44022.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PS enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051323 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other:

Alignment_scores:
Quality: 84.00 Length: 84
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x AA051323 ..

Align seg 1/1 to: AA051323 from: 1 to: 711

```

105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyr 121
|||||
301 ACCCCCTACCCATATGACAGAGGTTCTGCGTGAAGTGAATACCAATA 350
121 rSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluA 138
|||||
351 TTCTCGATATATGATGATGATGATGATGATGATGATGATGATGATGATG 400
138 rGluHisArgAsnArgGluTyrTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
401 GATTACATGCTAACAGGAAATATAGACCGGATATACAGAAATCTGAAT 450
155 lIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
|||||
451 ATAGCTCCGCGCAGAGGATGATGATGATGATGATGATGATGATGATGAT 500
171 sGlnAlaTyrPargGluGluProTyrPileHisHisAlaProGlnGlyCysG 188
|||||
501 CCAAGCTTGAGAGAGAACCTCGATTCATCATGCACCAAGGTTGTGG 550
188 lY 188
||
551 GA 552

```

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051324

seq_documentation_block:

ID AA051324 standard; DNA; 711 BP.

XX AA051324:

XX 08-DEC-1993 (first entry)

DE Encodes Ser-106 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 304..306

XX /tag= a

XX /note= "wild-type CCR(Pro) mutated to Acr(Ser)"

XX W09313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOGINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizaza M, Rappuoli R;

XX WPI; 1993-227320/28.

DR P-PSDB: AAR44023.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

The wild-type sequence coding for the A subunit of the heat labile toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AA051314-051326). The invention relates to immunogenic, detoxified LT-A proteins and their use in vaccines to protect against enterotoxigenic E.coli. Sequence AA051324 is a combination of the wild-type coding sequence and the mutagenic primer sequence used to introduce the preferred mutation. (Amino acid numbering is based on the cholera toxin A subunit sequence).

Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other:

Alignment_scores:
Quality: 82.00 Length: 82
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x AA051324 ..

Align seg 1/1 to: AA051324 from: 1 to: 711

```

107 HisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyrSerG 123
|||||
307 CACCCATATGACAGAGGTTCTGCGTGAAGTGAATACCAATATCTGA 356
123 nIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuH 140
|||||
357 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 406
140 lAsrArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
|||||
407 ATCTGACAGAGGAAATATAGACCGGATATACAGAAATCTGAATATAGCT 456
157 ProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAl 173
|||||
457 CCGGACAGAGGATGTTACAGATTAGCAGGTTTCCACCGGATACCAACG 506
173 aTrrPargGluGluProTyrPileHisHisAlaProGlnGlyCysGly 188
|||||
507 TTGGAGAGAGAAACCTCGATTCATCATGCACCAAGGTTGTGGGA 552

```

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051320

seq_documentation_block:

ID AA051320 standard; DNA; 711 BP.

XX AA051320:

XX 08-DEC-1993 (first entry)

DE Encodes Glu-107 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 307..309

XX /tag= a

XX /note= "wild-type CAC(His) mutated to GAG(Glu)"

XX M09313202-A.
 PN 08-JUL-1993.
 XX
 PD 30-DEC-1992; 92WO-EP03016.
 XX
 PR 31-DEC-1991; 91IT-0MI3513.
 XX
 PA (BIOC-) BIOCINE SCLAVO SPA.
 XX
 PI Domenighini M, Hol W, Piazza M, Rappuoli R;
 PI WPI: 1993-227320/28.
 DR P-PSDB: AAR44019.
 DR
 XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-Q51326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051320 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 234 A; 121 C; 158 G; 198 T; 0 other;

alignment_scores:
 Quality: 81.00 Length: 81
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051320 ..

Align seg 1/1 to: AA051320 from: 1 to: 711

108 PRTYRGLUGLNGLVAlserAlalaLeuGlyIlePRTYRserGlnI 124
 ||||||||||||||||||||||||||||||||||||||||||||
 310 CCATATGACACAGAGAGGTTCGTGAGTGAATACCATATTCGAGAT 359
 124 eTYRGLYTPYRArgValAsnPhelYValIleAspGluArgLeuH15A 141
 ||||||||||||||||||||||||||||||||||||||||||||
 360 AATATGATGATGATACGTAAATTTGTGTGATGATGAACGATTCATC 409
 141 rGAsnArgGluTYRArgAspArgTYRArgAsnLeuAsnIleAlaPro 157
 ||||||||||||||||||||||||||||||||||||||||||||
 410 GTTACAGGGAATATACAGACCGGTATACAGAAATCTGAATATAGCTCCG 459
 158 AlAGluAspGlyTYRArgLeuAlaGlyPheProProAspHisGlnAlaTr 174
 ||||||||||||||||||||||||||||||||||||||||||||
 460 GCAGAGAGATGGTTACAGATTAGCAGGTTCCACCGGATCACCACCACTTG 509
 174 pARGGluArgProTyrPheHisHisAlaProGlnGlyCysGly 188
 ||||||||||||||||||||||||||||||||||||||||||||
 510 GAGAGAGAAACCTGGATTCATCATCCACCAAGGTTGTGGA 552

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051325

seq_documentation_block:

ID AA051325 standard; DNA; 711 BP.

AC AA051325;

XX

DT 08-DEC-1993 (first entry)
 XX
 DE Encodes Glu-114 E.coli heat labile toxin subunit A.
 XX
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
 KW Protomer A; site-directed mutagenesis; reduced toxicity;
 KW ADP-ribosyltransferase activity; ss.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 328..330
 FT /*tag=a
 FT /note="wild-type TCT(Ser) mutated to GAA(Glu)"

XX M09313202-A.
 PN 08-JUL-1993.
 XX
 PD 30-DEC-1992; 92WO-EP03016.
 XX
 PR 31-DEC-1991; 91IT-0MI3513.
 XX
 PA (BIOC-) BIOCINE SCLAVO SPA.
 XX
 PI Domenighini M, Hol W, Piazza M, Rappuoli R;
 PI WPI: 1993-227320/28.
 DR P-PSDB: AAR44024.
 DR
 XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX

CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-Q51326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051325 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

alignment_scores:
 Quality: 74.00 Length: 74
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051325 ..

Align seg 1/1 to: AA051325 from: 1 to: 711

115 AlAlaGlyIleProTYRserGlnIleTYRGLYTPYRArgValAs 131
 ||||||||||||||||||||||||||||||||||||||||||||
 331 GCSTTAGGTGGAATACCATATTCACAGATATGAGATGATCGTTAA 360
 131 nPhelYValIleAspGluArgLeuHisArgAsnArgGluTYRArgAspA 148
 ||||||||||||||||||||||||||||||||||||||||||||
 381 TTTTGGTGATTCATGATGAACGATTACATCGTAACAGGAAATATAGAGACC 430
 148 rGTTYTYRArgAsnLeuAsnIleAlaProAlaGluAspGlyTYRArgLeu 164
 ||||||||||||||||||||||||||||||||||||||||||||
 431 GGTATATACAGAAATCTGAATATAGCTCCGACAGAGATGCTTACAGATTA 480
 165 AlAGlyPheProProAspHisGlnAlaTrpArgGluGluArgProTyrPheH 181

|||||
481 GCAGGTTTCCACCGGATCACCAGCTGGAGAGAGAACCCGTGATCA 530
181 SHSAlaProGInGlyCysGly 188
|||||
531 TCATGCAACCAAGGTTGTGGA 552

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51326

seq_documentation_block:
ID AAQ51326 standard; DNA; 711 BP.

XX AAQ51326;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-114 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS Key Location/Qualifiers

FT misc_difference 328..330

FT /tag= a /note= "wild-type TCR(Ser) mutated to AAA(Lys)"

XX MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OM13513.

XX (BIOC-) BIOGINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX WPI; 1993-227320/28.

XX P-PSDB; AAR44025.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

CC et al, J Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to

CC site-directed mutagenesis. Certain mutations were found to reduce

CC toxicity (see AAQ51314-051326). The invention relates to

CC immunogenic, detoxified LT-A proteins and their use in vaccines to

CC protect against enterotoxigenic E.coli. Sequence AAQ51326 is a

CC combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino

CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 237 A; 122 C; 156 G; 196 T; 0 other;

alignment_scores:

Quality: 74.00 Length: 74

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ51326 ..

Align seg 1/1 to: AAQ51326 from: 1 to: 711

115 AlaLeuGlyIleProTyrSerGlnIleTyrGlyTrpTyrArgValas 131
|||||
331 GCCTTAGGTGGATACCATTTCTCAGATATATGATGTGATGTGTA 380
131 npheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspa 148
|||||
381 TTTTGGTGTGATTTGATGACGATTTACATCTGTACAGGAAATATAGAGACC 430
148 rgtTyrTyrArgAsnLeuAsnIleAlaProIleGluAspGlyTyrArgLeu 164
|||||
431 GGTATTACAGAAATCTGAATATAGCTCCGGCAGAGATGGTTACAGATT 480
165 AlaGlyPheProAspHisGlnAlaTrpArgGluGluProTrpIleH 181
|||||
481 GCAGGTTTCCACCGGATCACCAGCTGGAGAGAGAACCCGTGATCA 530
531 TCATGCAACCAAGGTTGTGGA 552

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AA03448

seq_documentation_block:
ID AA03448 standard; DNA; 777 BP.

XX AA03448;

DT 02-MAY-1996 (first entry)

DE Shigella flexneri thermolabile enterotoxin gene fragment LTF.

XX Thermolabile enterotoxin; subunit-A; subunit-B; detection; probe;

KW LTF fragment; enterobacteria; ss.

XX Shigella flexneri.

XX RU2031948-C1.

XX 27-MAR-1995.

XX 11-DEC-1991; 91RU-0016860.

XX 11-DEC-1991; 91SU-4016860.

XX (NIZH-) NIZHEGOROD EPIDEMIOLOGY MICROBIOLOG INST.

XX Mazepa VN, Skobio LE, Ulanova TI;

XX WPI; 1995-343066/44.

XX LTF DNA fragment for recognition of thermo:labile enterobacterium

PT enterotoxin gene - can be used as a probe for recognition of

PT LT-toxin DNA gene

XX Claim 1; Columns 3-4; 5pp; Russian.

XX The present sequence is that of the LTF fragment which encodes part

CC of the thermolabile enterotoxin subunit-A sequence and all of the

CC subunit-B sequence of Shigella flexneri. The DNA fragment is useful

CC as a probe for specific detection of toxigenic enterobacteria.

XX Sequence 777 BP; 288 A; 141 C; 153 G; 195 T; 0 other;

alignment_scores:

Quality: 30.00 Length: 30

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA03448 ..

Align seg 1/1 to: AAT03448 from: 1 to: 777

173 AlarpragglugluProtrpIleHisAlaprogInglyCysGlyAs 189

3 GCTTGGAGAGAGAACCTGGATTTCATCATGACACCACAGGTTGGAAA 52

189 nSerSerArgThrIleThrGlyAspThrCysAsnGluGlu 202

53 TTCATCAAGAACATTCAGGTGATCTTGTAATGAGAGAG 92

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: AAN30099

seq_documentation_block:

ID AAN30099 standard; DNA; 795 BP.

AC AAN30099;

DT 02-JUN-1992 (first entry)

DE Sequence of HindII fragment from pEBK620 contg. an enterotoxoid gene.

XX Enterotoxoid; immunogen; vaccine; ds.

XX Key Location/Qualifiers

FT misc_feature 1..117

FT misc_feature 118..213

FT misc_feature 214..426

FT misc_feature 427..630

FT misc_feature 631

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

FT misc_feature 196 bp

alignment_scores:

Quality: 30.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

Alignment_block:

US-09-528-682-1 x AAN30099

Align seg 1/1 to: AAN30099 from: 1 to: 795

198 ThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214

694 ACTGTGATGAGAGAGACCCAGATCTGACCAATATATTCACAGGAGATA 743

214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227

744 TCAATCAAAAGCTTAAGCAGCATATTTTCAGACTATCAG 783

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAF74431

seq_documentation_block:

ID AAF74431 standard; DNA; 347 BP.

AC AAF74431;

DT 08-MAY-2001 (first entry)

DE Vibrio cholerae ctxA gene 347 bp region nucleotide sequence.

XX Vibrio cholerae: ctxA gene; detection; identification; toxigenic; ds.

XX Vibrio cholerae.

XX ZA200001349-A.

XX 27-DEC-2000.

XX 16-MAR-2000; 2000ZA-0001349.

XX 16-MAR-2000; 2000ZA-0001349.

XX (UYPR-) UNIV PRETORIA.

XX Theron J, Venter SN, Brozel VS, Du Preez M;

XX WPI; 2001-191829/19.

XX Novel oligonucleotide primer useful for PCR amplification of toxigenic

XX Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae

XX organisms present in the sample

XX Example 2; Fig 7; 44pp: English.

XX The present invention describes a method for amplifying (a) toxigenic

XX Vibrio cholerae nucleic acid by: (a) performing a polymerase chain

XX reaction (PCR); by combining a medium containing toxigenic V. cholerae

XX nucleic acid with an oligonucleotide primer (1) which will bind to a

XX section of V. cholerae nucleic acid, in an environment suitable for

XX carrying out PCR; (b) combining the medium with a second primer which

XX will bind downstream to the first primer to a section of the

XX complementary strand of the V. cholerae nucleic acid; and (c) amplifying

XX corresponding sections of V. cholerae nucleic acid present in the medium.

XX The method is useful for detecting toxigenic V. cholerae organisms or

XX toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental

XX water sample e.g., tap water sample, river water sample, lake water

CC sample or a sewage sample. The method is useful in PCR assays for a

CC detecting a microorganism or nucleic acid of the microorganism in a

CC sample. (1) and amplification products are useful for detecting toxigenic

CC V. cholerae nucleic acid in a sample. The PCR method involving (1) is

CC rapid, economical, specific and sensitive for detecting toxigenic

CC V. cholerae in water samples. The present sequence represents a

CC region from the V. cholerae ctxA gene, which is used in an example from

CC the present invention.
 XX
 SQ Sequence 347 BP; 101 A; 66 C; 77 G; 103 T; 0 other;

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAF74431 ..

Align seg 1/1 to: AAF74431 from: 1 to: 347

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 GAGTACTTTGACCGAGGACTCAATGAAATGCAACCTTATGATCATGC 116
 ||||||||||||||||||||||||||||||||||||||||||||||||
 45 AARGGLYThrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 117 AAGAGGAACTCAGACGGGATTGTAGG 144

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74430

seq_documentation_block:

ID AAF74430 standard; DNA: 564 BP.

AC AAF74430;

DT 08-MAY-2001 (first entry)

DE Vibrio cholerae ctxA gene 564 bp region nucleotide sequence.

KM Vibrio cholerae: ctxA gene; detection; identification; toxigenic; ds.

OS Vibrio cholerae.

XX ZA200001349-A.

PD 27-DEC-2000.

PE 16-MAR-2000; 2000ZA-0001349.

PR 16-MAR-2000; 2000ZA-0001349.

PA (UYPR-) UNIV PRETORIA.

PI Theron J, Venter SN, Brozel VS, Du Preez M;

DR WPI; 2001-191829/19.

PT Novel oligonucleotide primer useful for PCR amplification of toxigenic
 Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
 PT organisms present in the sample

PS Example 1; Fig 6; 44pp: English.

XX The present invention describes a method for amplifying (A) toxigenic
 CC Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
 CC reaction (PCR), by combining a medium containing toxigenic V. cholerae
 CC nucleic acid with an oligonucleotide primer (I) which will bind to a
 CC section of V. cholerae nucleic acid, in an environment suitable for
 CC carrying out PCR; (b) combining the medium with a second primer which
 CC will bind downstream to the first primer to a section of the
 CC complementary strand of the V. cholerae nucleic acid; and (c) amplifying
 CC corresponding sections of V. cholerae nucleic acid present in the medium.
 CC The method is useful for detecting toxigenic V. cholerae organisms or
 CC toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
 CC water sample e.g., tap water sample, river water sample, lake water
 CC sample or a sewage sample. The method is useful in PCR assays for
 CC detecting a microorganism or nucleic acid of the microorganism in a
 CC sample. (I) and amplification products are useful for detecting toxigenic

CC V. cholerae nucleic acid in a sample. The PCR method involving (I) is
 CC rapid, economical, specific and sensitive for detecting toxigenic
 CC V. cholerae in water samples. The present sequence represents a
 CC region from the V. cholerae ctxA gene, which is used in an example from
 CC the present invention.

SQ Sequence 564 BP; 159 A; 105 C; 139 G; 161 T; 0 other;

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAF74430 ..

Align seg 1/1 to: AAF74430 from: 1 to: 564

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 GAGTACTTTGACCGAGGACTCAATGAAATGCAACCTTATGATCATGC 116
 ||||||||||||||||||||||||||||||||||||||||||||||||
 45 AARGGLYThrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 117 AAGAGGAACTCAGACGGGATTGTAGG 144

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51328

seq_documentation_block:

ID AAQ51328 standard; DNA: 723 BP.

AC AAQ51328;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-63 cholera toxin subunit A.

KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

OS

XX Key Location/Qualifiers

FT misc_difference 187..189

FT /*tag= a

FT /note= "Wild-type TCA(Ser) mutated to AAG(Lys)"

XX MO9313202-A.

PD 08-JUL-1993.

PE 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-0MI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizza M, Rappuoli R;

DR WPI; 1993-227320/28.

PT P-PSDB; AAR44027.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX Claim 3; Fig 2 and Page 48; 60pp; English.
 XX The wild-type sequence coding for the A subunit of the cholera
 CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AAQ46318)
 CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051328 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SO Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051328 ..

Align seg 1/1 to: AA051328 from: 1 to: 723

29 GUTYTPheAspArgGlyThrGlnMetAsnIleAsnLeuTyraSPHSAI 45
|||||
85 GAGTACTTTCACCGAGTACTCAATGAAATATCAACCTTATGATCATGTC 134
45 AARGGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGCACTCAACGCGGATTTGTTAGG 162

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051329

seq_documentation_block:
ID AA051329 standard; DNA; 723 BP.

XX
AC AA051329;

DT 08-DEC-1993 (first entry)

XX
DE Encodes Lys-97 cholera toxin subunit A.

KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX
OS Vibrio cholerae.

XX
FH Key Location/Qualifiers

FT misc_difference 289..291

FT /tag= a

FT /note= "Wild-type GGT (Val) mutated to AAC(Lys)"

XX
PN WO9313202-A.

XX
PD 08-JUL-1993.

XX
PE 30-DEC-1992; 92WO-EP03016.

XX
PR 31-DEC-1991; 91IT-0M13513.

XX
PA (BIOC-) BIOCINE SCLAVO SPA.

XX
PI Domenighini M, Hol W, Piazza M, Rappuoli R;

XX
DR WPI: 1993-227320/28.

XX
DR P-PSDB: AAR44028.

XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX
PT - useful as vaccines against infection by Vibrio cholerae and
XX enterotoxin producing Escherichia coli
XX
PS Claim 3; Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
XX toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
XX was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051329 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SO Sequence 723 BP; 229 A; 125 C; 163 G; 206 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051329 ..

Align seg 1/1 to: AA051329 from: 1 to: 723

29 GUTYTPheAspArgGlyThrGlnMetAsnIleAsnLeuTyraSPHSAI 45
|||||
85 GAGTACTTTCACCGAGTACTCAATGAAATATCAACCTTATGATCATGTC 134
45 AARGGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGCACTCAACGCGGATTTGTTAGG 162

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051330

seq_documentation_block:
ID AA051330 standard; DNA; 723 BP.

XX
AC AA051330;

DT 08-DEC-1993 (first entry)

XX
DE Encodes Ser-106 cholera toxin subunit A.

KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX
OS Vibrio cholerae.

XX
FH Key Location/Qualifiers

FT misc_difference 316..318

FT /tag= a

FT /note= "Wild-type CCR(Pro) mutated to AGC(Ser)"

XX
PN WO9313202-A.

XX
PD 08-JUL-1993.

XX
PE 30-DEC-1992; 92WO-EP03016.

XX
PR 31-DEC-1991; 91IT-0M13513.

XX
PA (BIOC-) BIOCINE SCLAVO SPA.

XX
PI Domenighini M, Hol W, Piazza M, Rappuoli R;

XX
DR WPI: 1993-227320/28.

XX
DR P-PSDB: AAR44029.

XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX
PT - useful as vaccines against infection by Vibrio cholerae and
XX enterotoxin producing Escherichia coli
XX
PS Claim 3; Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
XX toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
XX was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051330 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051330 ..

Align seq 1/1 to: AA051330 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsn1LeasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCATCCTTTATGATCATGTC 134
|||||
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGCGGATTGTTAGG 162
```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051331

seq_documentation_block:
ID AA051331 standard; DNA; 723 BP.

AC AA051331;
XX
DT 08-DEC-1993 (first entry)

DE Encodes Asn-107 cholera toxin subunit A.

XX Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

XX
FH Key Location/Qualifiers
FT misc_difference 319..321

FT /*tag= a
/*note= "wild-type CAT(His) mutated to AAC(Asn)"

XX WO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizsa M, Rappuoli R.

XX WPI; 1993-227320/28.

DR P-PSDB; AAR44030.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051331 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 125 C; 163 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051331 ..

Align seq 1/1 to: AA051331 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsn1LeasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCATCCTTTATGATCATGTC 134
|||||
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGCGGATTGTTAGG 162
```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051332

seq_documentation_block:
ID AA051332 standard; DNA; 723 BP.

AC AA051332;
XX
DT 08-DEC-1993 (first entry)

DE Encodes Ser-110 cholera toxin subunit A.

XX Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

XX
FH Key Location/Qualifiers
FT misc_difference 328..330

FT /*tag= a
/*note= "wild-type GAA(Glu) mutated to TCG(Ser)"

XX WO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizsa M, Rappuoli R.

XX WPI; 1993-227320/28.

DR P-PSDB; AAR44031.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (Cr-A) (Mekalanos et al., Nature 306, 551 (1993) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (Ct_A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051334 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.

XX
SQ Sequence 723 BP; 229 A; 124 C; 164 G; 206 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051334 ..

Align seg 1/1 to: AA051334 from: 1 to: 723

29 G|U|T|Y|P|h|e|a|s|p|a|r|g|l|y|h|r|g|i|m|e|t|a|s|i|l|e|a|s|n|l|e|u|t|r|a|s|p|h|i|s|a|l 45
|||
85 G|A|G|T|A|C|T|T|G|A|C|C|G|A|G|T|A|C|T|C|A|A|T|G|A|T|A|T|C|A|C|C|T|T|A|T|G|A|T|C|A|T|C|G 134

45 a|a|r|g|g|l|y|h|r|g|i|h|t|h|r|g|l|y|h|p|h|e|v|a|l|a|r|g 54
|||
135 A|A|G|A|G|A|A|C|T|C|A|G|A|C|G|G|A|T|T|G|T|G|T|A|G|G 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA046318

seq_documentation_block:

ID AA046318 standard; DNA; 723 BP.

XX AC AA046318;

XX DT 08-DEC-1993 (first entry)

XX DE Cholera toxin subunit A coding sequence.

XX KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

XX KW protomer A; site-directed mutagenesis; reduced toxicity;

XX KW ADP-ribosyltransferase activity; ss.

XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers

XX FT mat_peptide 1..720

XX FT /tag= a

XX FT /note= "CT-A"

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX PS Disclosure; Fig 2; 60pp; English.

CC This sequence encodes the A subunit of the cholera toxin (CT-A)

CC of Vibrio cholerae. The sequence was published by Mekalanos et al.,

CC Nature 306, 551 (1983). Mutations at selected positions within this

CC sequence have been found to reduce toxicity (see AA051327-051334). The
CC invention relates to immunogenic, detoxified CT-A proteins and their
CC use in vaccines to protect against cholera.

XX
SQ Sequence 723 BP; 227 A; 125 C; 163 G; 208 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA046318 ..

Align seg 1/1 to: AA046318 from: 1 to: 723

29 G|U|T|Y|P|h|e|a|s|p|a|r|g|l|y|h|r|g|i|m|e|t|a|s|i|l|e|a|s|n|l|e|u|t|r|a|s|p|h|i|s|a|l 45
|||
85 G|A|G|T|A|C|T|T|G|A|C|C|G|A|G|T|A|C|T|C|A|A|T|G|A|T|A|T|C|A|C|C|T|T|A|T|G|A|T|C|A|T|C|G 134

45 a|a|r|g|g|l|y|h|r|g|i|h|t|h|r|g|l|y|h|p|h|e|v|a|l|a|r|g 54
|||
135 A|A|G|A|G|A|A|C|T|C|A|G|A|C|G|G|A|T|T|G|T|G|T|A|G|G 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV81596

seq_documentation_block:

ID AAV81596 standard; DNA; 723 BP.

XX AC AAV81596;

XX DT 11-MAY-1999 (first entry)

XX DE Cholera toxin gene.

XX KW A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;

XX KW parenteral adjuvant; antigen; antigen; immunisation; humoral response;

XX KW cell-mediated immune response; virus; bacterium; parasite; fungus;

XX KW tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;

XX KW systemic lupus erythematosus; Alzheimer's disease; diagnosis; ss.

XX OS Vibrio cholerae.

XX OS

XX PN WO9842375-A1.

XX PD 01-OCT-1998.

XX PF 19-MAR-1998; 98WO-US05454.

XX PR 18-MAR-1998; 98US-0044696.

XX PR 21-MAR-1997; 97US-0041227.

XX PA (CHIR) CHIRON CORP.

XX PI Barchfeld G, Del Giudice G, Rappuoli R;

XX PI WPI; 1999-070064/06.

XX PT

XX PT

XX PT

XX PT

XX PS Disclosure; Fig 1A-B; 51pp; English.

CC This sequence corresponds to the coding region for the A subunit of the
CC cholera toxin, an example of a bacterial ADP-ribosylating toxin. A
CC mutant detoxified form of this protein is used in a parenteral adjuvant
CC composition, which comprises the detoxified protein, at least one
CC selected antigen and optionally a pharmaceutically acceptable (optionally
CC topical) vehicle. The adjuvant composition can be administered
CC parenterally in conjunction with at least one antigen in methods to

CC Immunise vertebrate subjects. The adjuvant has the ability to enhance
CC the humoral and cell-mediated immune responses elicited by the antigen
CC (e.g. by making the antigen more strongly immunogenic or necessitating
CC fewer/lower antigen doses). It can be administered prior/subsequent to
CC the antigen, and is preferably administered within a short space of time
CC to the same site; it can also be administered in isolation from antigens
CC as a boost following systemic or mucosal antigen administration. Most
CC preferably, the adjuvant is co-administered with the antigen in the
CC compositions and a pharmaceutically acceptable carrier. The antigen may
CC be derived from viruses, bacteria, parasites and fungi or may be tumour
CC antigens, self-antigens and allergens. The compositions are therefore
CC useful in the treatment and prevention of e.g. viral diseases, allergic
CC manifestations, diseases caused by pathogens (e.g. bacteria or
CC parasites), AIDS, autoimmune diseases (e.g. systemic lupus
CC erythematosus), Alzheimer's disease and cancers. The adjuvant can also
CC be used to prepare antibodies against selected antigen(s), useful
CC e.g. for diagnostic purposes or for antigen purification.

SQ Sequence 723 BP; 227 A; 125 C; 163 G; 208 T; 0 other;

alignment_scores: Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAV81596 ..

Align seg 1/1 to: AAV81596 from: 1 to: 723

29 GluTyRPhesPargGlyThrGlnMetAsnIleAsnLeuTyRAspHisAl 45
|||||
85 GAGTACTTTGACGAGGTACTCAATGAAATATCAACCTTATGATCATGCG 134

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGGGAACCTCAGCGGCGATTGTGAGG 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51157

seq_documentation_block:
ID AAA51157 standard; cDNA: 777 BP.

XX
AC AAA51157;

XX
DT 26-SEP-2000 (first entry)

XX
DE Plant-optimized V. cholerae toxin A subunit coding sequence.

XX
KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KW adjuvant; anti-bacterial; ss.

XX
OS Vibrio cholerae.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT CDS 1..777 /*tag= a

FT /product= cholera_toxin_subunit_A
XX
XX

PN WO200037609-A2.

PD 29-JUN-2000.

XX
PF 22-DEC-1999; 99WO-US30747.

XX
PR 22-DEC-1998; 98US-0113507.

XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.

XX
PA (ARNT/) ARNTZEN C J.
XX

PI Mason HS, Arntzen CJ;

XX
XX WPI; 2000-442653/38.

DR P-PDB; AAY96653.

XX

PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals

XX
XX
XX Disclosure; Fig 4A; 103pp; English.

XX
PS This plant-codon optimized cDNA encodes a synthetic Vibrio cholerae
CC cholera toxin (CT) A subunit (CT-A). The sequence has no cryptic signal
CC sequence and no CG (potential methylation sites) sequences. Novel
CC polynucleotides encode a mutant Escherichia coli heat-labile toxin (LT)
CC subunit A (LT-A) polypeptide or a mutant V. cholerae cholera toxin (CT)
CC A subunit (CT-A) polypeptide, which have reduced enzyme activity as
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
CC of the codons is altered to a plant preferred codon. The polynucleotide
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
CC a CT B subunit (CT-B). The polynucleotides are useful for the
CC transformation of plant cells for the production of transgenic plants to
CC produce edible vaccines, especially oral vaccines in transgenic plants
CC for the prophylactic or therapeutic treatment against E. coli or
CC V. cholerae. The mutant polypeptides are also useful as adjuvants.

SQ Sequence 777 BP; 209 A; 174 C; 191 G; 203 T; 0 other;

alignment_scores: Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51157 ..

Align seg 1/1 to: AAA51157 from: 1 to: 777

29 GluTyRPhesPargGlyThrGlnMetAsnIleAsnLeuTyRAspHisAl 45
|||||
139 GAGTACTTTGACAGGGGACTCAGATGAACTCAACCTTATGACCATGCG 188

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAGCGGCGATTGTGAGG 216

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51158

seq_documentation_block:
ID AAA51158 standard; cDNA: 777 BP.

XX
AC AAA51158;

XX
DT 26-SEP-2000 (first entry)

XX
DE Plant-optimized mutant V. cholerae toxin A subunit K63 coding sequence.

XX
KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KW adjuvant; anti-bacterial; S63K; ss.

XX
OS Vibrio cholerae.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT CDS 1..777 /*tag= a

FT /product= mutant_cholera_toxin_subunit_A
XX
XX

PN WO200037609-A2.

PD 29-JUN-2000.

XX
XX

PF 22-DEC-1999; 99WO-US30747.
XX
PR 22-DEC-1998; 98US-0113507.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
PI Mason HS, Arntzen CJ;
XX
DR WPI; 2000-442653/38.
XX
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
PS Claim 15; Fig 4B; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a S63K
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
SQ Sequence 777 BP; 211 A; 172 C; 192 G; 202 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51158 ..

Align seg 1/1 to: AAA51158 from: 1 to: 777

29 GluYrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyraSPHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTGACATGCAACCTTTATGACCATGC 188
45 aARGGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAACTGATTTGTGAGG 216

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51159
seq_documentation_block:
ID AAA51159 standard; cDNA; 777 BP.
XX
AC AAA51159;
XX
XX 26-SEP-2000 (first entry)
XX
DE Plant-optimized mutant V. cholerae toxin A subunit R72 coding sequence.
XX
XX
KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KW adjuvant; anti-bacterial; A72R; ss.
XX
XX Vibrio cholerae.
OS Synthetic.
OS
XX
FH key Location/Qualifiers

FT CDS 1..777
FT /*tag= a
FT /product= mutant_cholera_toxin_subunit_A
XX
XX
XX MO200037609-A2.
XX
XX
XX 29-JUN-2000.
XX
XX
XX
PF 22-DEC-1999; 99WO-US30747.
XX
XX
PR 22-DEC-1998; 98US-0113507.
XX
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
XX
PI Mason HS, Arntzen CJ;
XX
XX
DR WPI; 2000-442653/38.
DR P-PSDB; AAY96654.
XX
XX
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
XX Claim 15; Fig 4C; 103pp; English.
XX
XX
CC This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a A72R
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
XX
SQ Sequence 777 BP; 210 A; 174 C; 191 G; 202 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51159 ..

Align seg 1/1 to: AAA51159 from: 1 to: 777

29 GluYrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyraSPHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTGACATGCAACCTTTATGACCATGC 188
45 aARGGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAACTGATTTGTGAGG 216

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51160
seq_documentation_block:
ID AAA51160 standard; cDNA; 777 BP.
XX
XX
AC AAA51160;
XX
XX 26-SEP-2000 (first entry)
XX


```
DE Plant-optimized mutant V. cholerae toxin A subunit G192 coding sequence.
XX
XX Heat-labile toxin: CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KM adjuvant; anti-bacterial; R192G; ss.
XX
XX Vibrio cholerae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..777
FT CDS /tag= a
FT /product= mutant_cholera_toxin_subunit_A
XX
PN MO200037609-A2.
PD 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US30747.
XX
XX 22-DEC-1998; 98US-0113507.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
PI Mason HS, Arntzen CJ;
XX
XX WPI: 2000-442653/38.
DR P-PSDB; AAY96656.
XX
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX Claim 15: Fig 4D; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a R192G
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
SQ Sequence 777 BP; 208 A; 174 C; 191 G; 203 T; 1 U; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAS1160 ..
Align seg 1/1 to: AAS1160 from: 1 to: 777

29 GIUTYrPheasPARgLyThrGlnMeTAsnIleasNleuTYrAspHSAI 45
|||||
139 GAGTACTTGTGACAGGGGTACTGACATGCAACCTTTATGACCATGC 188
|||||
45 aARGGLyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACATGATTTGTGAGG 216
```

```
seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: AAN30169
seq_documentation_block:
ID AAN30169 standard; DNA; 1148 BP.
XX
XX AAN30169;
AC
XX 25-APR-1992 (first entry)
DT
XX
XX Sequence of PRIT 18014 encoding sub-unit A and B of cholera toxin.
DE
XX Vaccine; cholera; subunit A; subunit B; ss.
KM
XX
XX Vibrio cholerae.
OS
XX
FH Key Location/Qualifiers
FT misc_feature 1..777
FT /tag= a
FT /label= F1
FT /note= "Claimed fragment"
FT misc_feature 774..1148
FT /tag= b
FT /label= F2
FT /note= "Claimed fragment"
XX
XX EP95452-A.
PN
XX
XX 30-NOV-1983.
PD
XX
XX 23-MAY-1983; 83EP-0091416.
PF
XX
XX 24-MAY-1982; 82US-0381083.
PR
XX
XX (SMIR ) SMITH KLINE-RIT.
PA
PI Harford N, Dewilde M;
XX
XX WPI: 1983-834665/49.
DR
XX
XX DNA sequence coding for cholera toxin sub units - for
PT transforming host cells, useful in vaccine prodn.
PT
XX
XX Claim 2: Page 27-28; 46pp; French.
XX
XX The inventors claim a sequence contg. at least one fragment coding
CC for all or part of sub-units A and B of cholera toxin. The entire
CC sequence and two fragments (F1 and F2) are listed in the
CC specification. Also new are recombinant DNA molecules contg. these
CC sequences (Spec. PRIT 18014 contg. the whole sequence; PRIT 18041
CC contg. F1 and PRIT 10810 contg. F2) and transformed host cells (Spec.
CC E. coli ATCC 39052 contg. the whole sequence, ATCC 39053 contg. F1
CC and ATCC 39051 contg. F2).
XX
SQ Sequence 1148 BP; 382 A; 184 C; 235 G; 347 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAN30169 ..
Align seg 1/1 to: AAN30169 from: 1 to: 1148

29 GIUTYrPheasPARgLyThrGlnMeTAsnIleasNleuTYrAspHSAI 45
|||||
139 GAGTACTTGTGACAGGAGTACTCAATGATATCAACCTTTATGATCATGC 188
|||||
45 aARGGLyThrGlnThrGlyPheValArg 54
|||||
189 AAGGAACTCAGACGGGATTTGTAGG 216
```


seq_name: /SIDSS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA167270

seq_documentation_block:

ID AA167270 standard; DNA; 2056 BP.

AA167270;

11-FEB-2002 (first entry)

Cholera toxin A and B subunits encoding DNA sequence.

Enterotoxin; mutation; chimeric; immunostimulant; immune response; cholera toxin; ds.

Vibrio cholerae.

Location/Qualifiers

1..550

5'UTR

CDS

551..1327

/*tag- b

/product- "Cholera toxin A subunit"

/note- "AAG65991"

1324..1698

/*tag- c

/product- "Cholera toxin B subunit"

/note- "AAG65992"

/product- "B chain"

1699..2056

3'UTR

/*tag- d

WO200170257-A1.

27-SEP-2001.

16-MAR-2001; 2001MO-US08582.

17-MAR-2000; 2000US-190058P.

(UABR-) UAB RES FOUND.

Mcghee J, Kiyono H, Takeda Y, Ohmura M, Yamamoto S;

WPI: 2001-648368/74.

P-SDB; AAG65991, AAG65992.

Chimeric molecule useful as adjuvant for cell-mediated/humoral immunity

comprises first mutated A subunit of first enterotoxin and second

non-mutated subunit from second enterotoxin different from natural

enterotoxin

Disclosure; Fig 4; 23pp; English.

The invention provides a chimeric molecule that comprises a first subunit

which is mutated A subunit of first enterotoxin and a second non-mutated

subunit from a second enterotoxin which is different from the natural

enterotoxin which has been mutated to provide A subunit. A composition of

matter comprising the chimeric molecule in a pharmaceutically acceptable

carrier is useful for obtaining enhanced immune response of an organism

to an antigen by administering it to the antigen. The chimeric molecule

is useful to provide specific immune response to a particular

enterotoxin, as adjuvants for use with unrelated vaccines, and to

customize adjuvants to direct production of cell-mediated or humoral

immune responses. It is also useful as mucosal adjuvant for cell-mediated

or humoral immunity. The present sequence represents the DNA sequence of

Cholera toxin A and B subunits.

Sequence 2056 BP; 603 A; 327 C; 479 G; 647 T; 0 other;

alignment_scores: 26.00 Length: 26

Quality: 26.00

26

26

26

26

26

26

26

26

26

26

26

26

26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA167270 ..

Align seq 1/1 to: AA167270 from: 1 to: 2056

29 GluTyrPheAspArgGlyThrGlnMetAsn1leAsnLeuTyrAspHisAl 45

|||||

689 GAGTACTTGGACCGAGGTACTCAATATCATATCACTTATATCATATGC 78

|||||

45 AArgGlyThrGlnThrGlyPheValArg 54

|||||

739 AAGAGAACTCAGACGCGATTGTAGG 766

seq_name: /SIDSS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABA92644

seq_documentation_block:

ID ABA92644 standard; DNA; 7586 BP.

ABA92644;

25-MAR-2002 (first entry)

Cholera toxin A1 subunit sequence SEQ ID NO:1.

Cholera toxin; bacterial bleb; vaccine; eukaryotic expression cassette;

gene therapy; ds.

Vibrio cholerae.

WO200189535-A1.

29-NOV-2001.

24-MAY-2001; 2001MO-US16904.

24-MAY-2000; 2000US-206994P.

(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.

Powell RJ, Hone D;

WPI: 2002-122002/16.

Introducing and expressing a gene in animal cells, useful for

delivering a therapeutic agent to the cells, comprises infecting animal

cells with bacterial blebs containing a eukaryotic expression cassette

encoding the gene

Example 5; Fig 3; 83pp; English.

The present invention describes a method for introducing and expressing

a gene in animal cells, comprising infecting animal cells with bacterial

blebs containing a eukaryotic expression cassette encoding the gene.

The present invention also describes: (1) a bacterial bleb containing a

eukaryotic expression cassette; (2) a transformed animal cell including

a bacterial bleb, or a transfected product of the bacterial bleb;

(3) a method of delivering a eukaryotic expression cassette to an

animal cell, comprising mucosal transfection; and (4) a method of

polymerically treating an animal with a therapeutic agent encoded by a

polynucleotide coding sequence, by transfecting the animal with a

bacterial bleb containing a eukaryotic expression cassette including the

polynucleotide coding sequence. The method can be used for introducing

endogenous or foreign genes into animal cells using bacterial blebs as

vectors. The method allows for the delivery of eukaryotic expression

cassettes encoding the endogenous or foreign genes into animal cells or

tissues, and is useful for expressing e.g. vaccine antigens, gene

therapeutic agents, immunoregulatory agents, antisense RNAs, and

catalytic RNAs in animal cells or tissues. The bacterial blebs

containing the eukaryotic expression cassette can also be used to treat

animal cells cultured in vitro. The method can be used in gene therapy.

CC The present sequence represents the cholera toxin A1 subunit sequence,
 CC which is used in an example from the present invention for the production
 CC of non-pyrogenic blebs containing a DNA vaccine.

XX
 SQ Sequence 7586 BP; 1619 A; 1747 C; 1733 G; 1657 T; 830 other;

alignment_scores:

Quality:	26.00	Length:	26
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x ABA92644 ..

Align seg 1/1 to: ABA92644 from: 1 to: 7586

```

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
87 GAGTACTTACCGAGGTACTCAATGATATTCACCTTATGATCATGCG 136
45 AARGlyThrGlnThrGlyPheValArg 54
|||||
137 AAGAGCAACTCAGACGCGATTGTGTAGC 164

```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ12710

seq_documentation_block:
 ID AAQ12710 standard; DNA; 780 BP.

XX
 AC AAQ12710;

DT 08-OCT-1991 (first entry)

XX
 DE Cholera toxin A1 fragment.

XX Hybrid: fusion: membrane translocation: binding region: HIV;
 KM infection: toxin: steroid: hormone: monoclonal antibody; antigen;
 KM diphtheria; exotoxin; phenylketonuria; cholera; interleukin; IL-2;
 KM protease; epidermal growth factor; ricin; tetanus; hexosaminidase;
 KM Shiga-like toxin A; ligand; insulin; nuclease; ss.

XX
 OS Vibrio cholerae.

Key	Location/Qualifiers
FT CDS	1..777
FT	/*tag= a
FT	/label= CT-A1

PN WO9109671-A.

PD 11-JUL-1991.

PF 21-DEC-1990; 90WO-US07619.

PR 14-JUN-1990; 90US-0538276.

PR 22-DEC-1989; 89US-0456095.

PA (SERA-) SERAGEN INC.

PI Murphy JR;

DR WPI: 1991-222845/30.

DR P-PSDB: AARI3117.

XX Hybrid molecules for targeting chemical entity to cell - have
 PT membrane trans-locating and cell binding-regions and used to
 PT treat HIV infection, genetic enzyme-deficiency disorders etc.

PS Disclosure; Fig 5(1-2); 59pp; English.

CC Hybrid molecules are produced by covalently linking
 CC (1) a portion (A) of the binding domain of a cell-binding ligand,

CC allowing binding of the mol. to an animal cell;

CC (2) a portion (B) of a translocation domain of a protein able to
 CC translocate (C) across the cell cytoplasmic membrane, and the

CC and (3) a portion (C) which is to be introduced into the cell.

CC (A) is derived from a steroid or polypeptide hormone, a single-chain
 CC analogue of a monoclonal antibody able to bind an antigen expressed
 CC on the cell surface, or a polypeptide toxin.

CC (B) is derived from a toxin (e.g. diphtheria toxin or Pseudomonas
 CC exotoxin A).

CC (A) may be derived from insulin, interleukins 2, 3 or 6 or
 CC epidermal growth factor.

CC Suitable enzymes in (C) include cholera toxin (encoded by this
 CC sequence), ricin, tetanus toxin, hexosaminidase A, protease,
 CC nuclease, etc.

CC Specified examples are CT-A/DT-B'/IL-2, SLTA/DT-B'/IL-2,
 CC ricin A/DT-B'/IL-2, HIV-BP/DT-B'/IL-2 and the phenylalanine

CC hydroxylase-DT-B' or their biologically active mutants.

CC (CT-A= cholera toxin, DT-B'= truncated diphtheria toxin,
 CC SLTA= Shiga-like toxin A; HIV-BP= HIV protease binding protein.
 CC See also AAQ12710-12.

XX
 SQ Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;

alignment_scores:

Quality:	25.00	Length:	25
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AAQ12710 ..

Align seg 1/1 to: AAQ12710 from: 1 to: 780

```

30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaAr 46
|||||
142 TACTTTGACCGAGGTACTCAATGATATTCACCTTATGATCATGCAAG 191
46 GGLyThrGlnThrGlyPheValArg 54
|||||
192 AGGAACCTCAGACGCGATTGTGTAGC 216

```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAI91636

seq_documentation_block:
 ID AAT91636 standard; DNA; 780 BP.

XX
 AC AAT91636;

DT 27-MAR-1998 (first entry)

XX
 DE Vibrio cholerae toxin gene.

XX Cholera toxin A1; hybrid protein; cell delivery;
 KM cell binding ligand; translocation domain; diphtheria toxin B';
 KM interleukin-2; leukaemia; therapy; ss.

XX
 OS Vibrio cholerae.

Key	Location/Qualifiers
FT CDS	1..777
FT	/*tag= a
FT	/transl_except= (pos:358..360, aa:Ala)
FT	/transl_except= (pos:616..618, aa:Ala)

PN US5668255-A.

PD 16-SEP-1997.

PF 07-JUN-1984; 84US-0618199.

PR 27-JUN-1991; 91US-0722484.

PR 07-JUN-1984; 84US-0618199.


```
PR 25-APR-1985; 85US-0726808.
PR 07-JUN-1985; 85US-0742554.
PR 22-DEC-1989; 89US-0456095.
PR 14-JUN-1990; 90US-0538276.
PR 04-AUG-1993; 93US-0102387.
XX
XX (SERA-) SERAGEN INC.
XX
XX Murphy JR;
XX
XX WPI: 1997-470103/43.
XX
XX P-PSDB; AAM25785.
XX
XX New hybrid molecules for delivery of agents to cells - comprise a
XX binding domain of a cell binding ligand and a portion of a
XX translocation domain of a protein
XX
XX Example 2; Fig 5A-B; 30pp; English.
XX
XX This sequence comprises the Vibrio cholerae gene that encodes
XX cholera toxin (see AAM25785). DNA encoding an enzymatically active
XX A1 fragment of the cholera toxin was used to construct a cholera
XX toxin A1-diphtheria toxin B'-interleukin-2 (CTA-DTB-IL2) gene that
XX was expressed in E. coli. The hybrid protein can be isolated and
XX used e.g. as an adjunct to treatment with DTB-IL2 hybrid, which
XX targets the IL2 receptor, such as certain leukaemic T cells.
XX bearing the IL2 receptor, such as certain leukaemic T cells.
XX Prior application of the CTA-DTB-IL2 hybrid alleviates the
XX problem of endogenous IL2 competing with DTB-IL2 hybrid for IL2
XX receptors on T cells. Claimed hybrid proteins comprise a
XX translocation domain and a cell binding domain from e.g. a hormone,
XX growth factor or polypeptide toxin. The hybrid molecules can be
XX used for the delivery of agents (e.g. therapeutic genes, toxins,
XX detectable labels) into cells. The use of a translocation
XX mechanism ensures that the hybrid will be effective in relatively
XX low doses, since a high proportion of the substance of interest
XX will be taken into the targeted cells. The hybrid molecules can be
XX manufactured as a single hybrid recombinant protein, permitting
XX reproducibility, consistency, and the precise control of
XX composition.
XX
XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;
SQ
alignment_scores:
  Quality: 25.00 Length: 25
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AAT91636 ..
Align seg 1/1 to: AAT91636 from: 1 to: 780
30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaIar 46
|||||
142 TACTTTGACCGAGGCTACTCAATGATATCAACCTTTATGATCATGCAAG 191
|||||
46 GGLYThrglnThrGlyPheValArg 54
|||||
192 AGGAACCTCAGACGGGATTTGTTAGG 216
seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ30661
seq_documentation_block:
ID AAZ30661 standard; DNA: 780 BP.
XX
XX AAZ30661;
XX
XX 15-FEB-2000 (first entry)
XX
XX Vibrio cholerae A1 fragment toxin gene.
XX
```

```
KW Recombinant: hybrid; binding domain; ligand; animal cell; diphtheria;
KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;
KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;
KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;
KW adipocyte; cancer; virus; infection; antibody; ss.
XX
XX Vibrio cholerae.
XX
XX US5965406-A.
XX
XX 12-OCT-1999.
XX
XX 07-JUN-1995; 95US-0488246.
XX
XX 04-AUG-1993; 93US-0102387.
XX 07-JUN-1984; 84US-0618199.
XX 27-JUN-1991; 91US-07222484.
XX 25-APR-1985; 85US-0726808.
XX 07-JUN-1985; 85US-0742554.
XX 22-DEC-1989; 89US-0456095.
XX 14-JUN-1990; 90US-0538276.
XX
XX (SERA-) SERAGEN INC.
XX
XX Murphy JR;
XX
XX WPI: 1999-632431/54.
XX P-PSDB; AAY55890.
XX
XX Recombinant DNA molecule encoding a three part hybrid protein used in
XX the treatment of Aids and genetic deficiency diseases -
XX
XX Example 2; Fig 5; 31pp; English.
XX
XX The invention relates to a recombinant DNA molecule encoding a hybrid
XX protein comprising three parts: (a) the first part comprises a portion
XX of the binding domain of a cell-binding polypeptide ligand allowing the
XX hybrid protein to bind to an animal cell; (b) the second part comprises
XX a portion of a translocation domain of a naturally occurring protein
XX selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
XX toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
XX toxin, which translocate the third part of the across the cytoplasmic
XX membrane into the cytosol of the cell; and (c) the third part comprises
XX a polypeptide entity to be introduced into the cell, which is non-native
XX to the naturally occurring protein of (b). This sequence represents the
XX Vibrio cholerae toxin A1 fragment gene for use in generating the hybrid
XX of the invention. The hybrid molecule enables the direction of
XX appropriate therapy to affected cells, allowing them to function properly
XX and alleviate or cure the disease. The hybrid is especially used in
XX treating genetic deficiency diseases, by delivering to affected cells
XX an enzyme supplying the missing function, to supplementing cellular
XX levels of a particular enzyme or a scarce precursor or cofactor, to
XX directing toxins or other poisons to destroy particular cells (such as
XX adipocytes, cancer cell, or virus infected-cells), to counteracting viral
XX infections such as HIV, by introducing appropriate antibodies to viral
XX proteins. It is also involved in the process of getting non-therapeutic
XX substances such as detectable labels into cells.
XX
XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;
SQ
alignment_scores:
  Quality: 25.00 Length: 25
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AAZ30661 ..
Align seg 1/1 to: AAZ30661 from: 1 to: 780
30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaIar 46
|||||
```


142 TACTTTGACCGAGTCAATGATATCAACCTTATGATCATGCAG 191
 46 ggGlyThrGlnThrGlyPheValArg 54
 ||||||||||||||||||
 192 AGGAACTCAGACGCGGATTTGTTAGG 216

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA2000.DAT:AA290017

seq_documentation_block:

ID AA290017 standard; DNA; 780 BP.

AA290017:

05-MAY-2000 (first entry)

Vibrio cholerae toxin gene.

Cholera toxin; hybrid protein; translocation domain; cell binding domain;
 genetic deficiency disease; cell targeting; cell destruction; cancer;
 adipocyte; enzyme delivery; anti-viral; HIV; ds.

Vibrio cholerae.

US6022950-A.

08-FEB-2000.

07-JUN-1995; 95US-0479510.

07-JUN-1984; 84US-0618159.

27-JUN-1991; 91US-0722484.

25-APR-1985; 85US-0726808.

07-JUN-1985; 85US-0742554.

22-DEC-1989; 89US-0456095.

14-JUN-1990; 90US-0538276.

04-AUG-1993; 93US-0102387.

(SERA-) SERAGEN INC.

Murphy JR;

WPI; 2000-160390/14.

P-PSDB; AAY78590.

Example 2; Fig 5; 32pp; English.

This sequence represents the Vibrio cholerae toxin gene. The toxin can be included in the hybrid protein of the invention and used to destroy or modify the cell that the hybrid protein is targeted to. The hybrid protein comprises a first part which is a portion of the binding domain of a cell-binding ligand, effective to cause the hybrid molecule to bind to a cell of an animal. The second part comprises a portion of a translocation domain of a naturally occurring protein (e.g. the translocation domain of diphtheria toxin) the second part translocates the third part across the cytoplasmic membrane and into the cytosol of the cell. The third part comprises a chemical entity to be introduced into the cell, where each of the first and third part is non-native with respect to naturally occurring protein. The covalent bond attaching the second and third part is cleavable. The toxin encoded by the present sequence can form part of the third portion of the hybrid protein. The cell binding domain binds to a specific cell and the translocation domain transfers the hybrid molecule across the cell membrane into the cytosol. The third part of the protein, linked to the translocation domain through a cleavable bond, can then carry out its function. The hybrid molecules are useful for treating genetic deficiency diseases by delivering to affected cells an enzyme supplying the missing function, to supplement cellular levels of a particular enzyme or a scarce precursor or cofactor, to direct toxins or other poisons to destroy particular cells (such as adipocytes, cancer cells, or virus-infected cells), and to counteract

CC viral infections such as HIV by introducing into appropriate cells
 CC antibodies to viral proteins.
 CC XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;

alignment_scores:

Quality:	25.00	Length:	25
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AA290017

Align seg 1/1 to: AA290017 from: 1 to: 780

30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaAr 46
 ||||||||||||||||||
 142 TACTTTGACCGAGTCAATGATATCAACCTTATGATCATGCAG 191
 46 ggGlyThrGlnThrGlyPheValArg 54
 ||||||||||||||||||
 192 AGGAACTCAGACGCGGATTTGTTAGG 216

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051327

seq_documentation_block:

ID AA051327 standard; DNA; 723 BP.

AA051327:

08-DEC-1993 (first entry)

Encodes Asp-53 cholera toxin subunit A.

Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

protomer A; site-directed mutagenesis; reduced toxicity;

ADP-ribosyltransferase activity; ss.

Vibrio cholerae.

Key Location/Qualifiers

misc_difference 156..158

W09313202-A.

08-JUL-1993.

30-DEC-1992; 92WO-EP03016.

31-DEC-1991; 91IT-OM13513.

(BIOC-) BIOCINE SCLAVO SPA.

Domenighini M, Hol W, Piazza M, Rappuoli R;

WPI; 1993-227320/28.

P-PSDB; AAR44026.

Immunogenic detoxified mutant cholera toxin and heat labile toxin

- useful as vaccines against infection by Vibrio cholerae and

enterotoxin producing Escherichia coli

Claim 3; Fig 2 and Page 48; 60pp; English.

The wild-type sequence coding for the A subunit of the cholera toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AA051327-051334). The invention relates to immunogenic, detoxified CT-A proteins and their use in vaccines to protect against Vibrio cholerae. Sequence AA051327 is a

CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 126 C; 163 G; 206 T; 0 other;

alignment_scores:
Quality: 24.00 Length: 24
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ31327 ..

Align seg 1/1 to: AAQ31327 from: 1 to: 723

29 GluTyrPheAspArgGlyThrGlnMetAsnLeuTyrrAspHisAl 45
|||||
85 GAGTACTTGTGACGAGTACTCAATGATGATCAACCTTATGATCATGC 134
45 AARGlyThrGlnThrGlyPhe 52
|||||
135 AAGAGGAACTCAGACGCGATT 156

seq_name: /SID55/gcgdata/geneseq/geneseqn.embl/NA1992.DAT:AAQ31263

seq_documentation_block:

ID AAQ31263 standard: DNA; 777 BP.

XX AAQ31263;

XX 25-MAR-1993 (first entry)

XX Mutant A subunit of CT.

XX Cholera toxin; A1; A2; B; subunit; mutation; ss.

XX Vibrio cholerae.

XX Key Location/Qualifiers

FT CDS 1..777

FT sig_peptide 1..34

FT mat_peptide 55..774

FT misc_feature 630

FT misc_feature 636

FT misc_feature 637..777

FT misc_feature 73..75

FT misc_difference 79..81

FT misc_difference 85..87

FT misc_difference 184..186

FT misc_difference 262..264

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT misc_difference 388..390

FT /tag= m
FT /note= "Glu112 codon mutation, see CC"
FT misc_feature 588
FT /tag= n
FT /note= "start of C-terminal truncation, see CC"

WO9219265-A.
12-NOV-1992.

04-MAY-1992; 92WO-US03703.

02-MAY-1991; 91US-0694733.

(AMGE-) AMGEN.
(UYSC-) UNIV SOUTHERN CALIFORNIA.

Burnette WN, Kaslow HR;

WPI; 1992-398532/48.

P-PSDB; AAR28830.

Recombinant DNA encoding analogue of Vibrio cholerae catalytic

sub-unit - useful as a vaccine for preventing cholera, has little

or no catalytic activity

Claim 7-8; Page 49-50 + Fig 1A; 74pp; English.

A recombinant mol. encoding an analogue of the catalytic subunit of
cholera toxin (CT), where the analogue has reduced or no catalytic
activity associated with CT. reactogenically is an analogue of the A
region, esp. A1 region, comprising a site-specific mutation in the
CC region bounded by the codons for Met1 and Arg192 or Ser194, esp. one
or more of Arg7, Arg11, Asp9, His44, His70 and Glu112 (see tag h-m
of Features Table). The analogue may also comprise a truncation of
the C-terminal portion starting at Trp179 (see tag n of Features Table).
The DNA may also encode subunit B of CT, esp. the B oligomer in
native form or which has been genetically engineered (see AAQ31264).

Sequence 777 BP; 235 A; 125 C; 165 G; 234 T; 18 other;

alignment_scores:
Quality: 22.00 Length: 22
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ31263 ..

Align seg 1/1 to: AAQ31263 from: 1 to: 777

81 SerThrTyrTrpIleTyrValIleAlaThrAlaProAsnMetPheAsn 97
|||||
295 TCTACTTATTATATATATGTTATAGCCACTGACCCAACTGTTAAAG 344
97 IAsnAspValLeuGly 102
|||||
345 TAATGATGATATTAGG 360

gb_gss:BH268086	8.00	124.52	1.1e+03	539	BH268086 CH230-56a10.TJ CHORI-2	gb_gss:BH249846	8.00	121.70	1.6e+03	764	BH249846 BCGAM63TF BCGA Bras
gb_est1:AV882538	8.00	124.51	1.1e+03	540	AV882538 AV882538 Nori Satoh un	gb_est2:BI658846	8.00	121.66	1.6e+03	768	BI658846 602328613F1 NIH_MGC
gb_est2:BM134967	8.00	124.46	1.1e+03	543	BM134967 MHED040.C12.E23S Whee	gb_est2:BG446656	8.00	121.52	1.6e+03	781	BG446656 602420826F1 NIH_MGC
gb_gss:AA2159324	8.00	124.45	1.1e+03	543	AA215932 H. brucei shepard genc	gb_est2:BG789155	8.00	121.51	1.6e+03	782	BG789155 SPBAM00911F1 See ur
gb_est2:BF26615	8.00	124.43	1.1e+03	545	BF26615 HV_CBE0014E20F Hordeum	gb_gss:AO687372	8.00	121.44	1.6e+03	789	AO687372 nxbx0074119F CUCI R
gb_gss:015353	8.00	124.35	1.1e+03	551	AJ227316 Clona intestinalis get	gb_est1:BE033932	8.00	121.43	1.6e+03	790	BE033932 MG020C06 MG Mesembry
gb_est1:AW755146	8.00	124.30	1.1e+03	551	AJ227316 Clona intestinalis get	gb_gss:ONS03A1W	8.00	121.43	1.6e+03	790	AL234653 Tetradodon nigrovir
gb_gss:BH338148	8.00	124.30	1.1e+03	554	BC62819 MHED0945.F06.K11S Whee	gb_est1:AV897192	8.00	121.30	1.7e+03	803	AV897192 AV897192 Nori Satoh
gb_gss:BH338148	8.00	124.30	1.1e+03	554	BC62819 MHED0945.F06.K11S Whee	gb_gss:BH465346	8.00	121.27	1.7e+03	805	BH465346 BCGY044TF BCGY Bras
gb_est1:AL510917	8.00	124.24	1.1e+03	558	AL510917 CH230-12412.TVB CHORI-	gb_est1:AV681452	8.00	121.25	1.7e+03	807	AV681452 AV681452 GKA Homo s
gb_est2:BM178462	8.00	124.21	1.1e+03	560	AL510917 CH230-12412.TVB CHORI-	gb_gss:AO745347	8.00	121.17	1.7e+03	816	AO745347 HS_2276.AL.B11.T7C
gb_gss:AZ709793	8.00	124.16	1.1e+03	564	AL510917 CH230-12412.TVB CHORI-	gb_est2:BG369434	8.00	121.15	1.7e+03	818	BG369434 HYSME0024G08F Hord
gb_est1:AA733970	8.00	124.13	1.1e+03	566	AZ709793 RPCI-24-116118.TJ RPCI	gb_est2:BI2556096	8.00	121.14	1.7e+03	819	BI2556096 602978931F1 NCI_CGA
gb_est2:BG414316	8.00	124.13	1.1e+03	566	AA733970 467.P1F62 Plnus tede	gb_est2:BI956846	8.00	121.14	1.7e+03	819	BI956846 HYSME0005K09F Hord
gb_est1:AV933568	8.00	124.06	1.2e+03	571	BG414316 HYSME0001G05F Hordeum	gb_gss:CMS02CA6	8.00	121.05	1.7e+03	828	BI956846 HYSME0005K09F Hord
gb_est2:BG414316	8.00	124.06	1.2e+03	571	AV933568 AV933568 K. Sato unpu	gb_gss:CMS03A3W	8.00	121.04	1.7e+03	829	AL264353 Tetradodon nigrovir
gb_est2:BR483656	8.00	124.00	1.2e+03	575	BG414316 HYSME0001E09F Hordeum	gb_est1:BE194447	8.00	120.84	1.7e+03	844	BE194407 HYSME0085H01F Hord
gb_est2:BF065521	8.00	123.97	1.2e+03	577	BF483656 WHE2336.C12.F24ZS Whee	gb_est2:BM018143	8.00	120.86	1.7e+03	847	BM018143 603645733F1 NIH_MGC
gb_est1:AL510920	8.00	123.96	1.2e+03	578	BF065521 HV_CBE0012K16F Hordeum	gb_gss:CMS04S0U	8.00	120.80	1.8e+03	854	AL305679 Tetradodon nigrovir
gb_gss:AO2313446	8.00	123.90	1.2e+03	582	AL510920 AL510920 Hordeum vulg	gb_est2:BF575094	8.00	120.76	1.8e+03	858	BF575094 F27A9-SP6.1 IGF Arabi
gb_gss:AO2313446	8.00	123.87	1.2e+03	584	AO2313446 RPCI11-101B1.TJ RPCI-1	gb_est2:BF120183	8.00	120.65	1.8e+03	870	BF120183 602134766F1 NCI_CGA
gb_est2:BI640414	8.00	123.83	1.2e+03	587	AO274058 nxbx0032118F CUCI Rice	gb_est2:BF575766	8.00	120.52	1.8e+03	884	BF575766 602135396F1 NIH_MGC
gb_est1:AM014756	8.00	123.82	1.2e+03	588	BI640414 SD23280.5Prime SD Dros	gb_est2:BI222903	8.00	120.47	1.8e+03	889	BI222903 602941152F1 NIH_MGC
gb_gss:AO2156267	8.00	123.64	1.2e+03	594	AM014756 UT-H-B10-aae-e-11-0-UT	gb_gss:CMS04HK9	8.00	120.42	1.9e+03	917	AL291942 Tetradodon nigrovir
gb_est1:AV9336810	8.00	123.70	1.2e+03	598	BH260927 CH230-90H24.TV CHORI-2	gb_gss:CMS02FC6N	8.00	120.22	1.9e+03	949	AL1334815 Tetradodon nigrovir
gb_est2:BI074785	8.00	123.68	1.2e+03	598	AV9336810 AV9336810 K. Sato unpu	gb_gss:CMS03A3W	8.00	119.94	2.0e+03	959	AL303456 Tetradodon nigrovir
gb_est1:AV9336810	8.00	123.64	1.2e+03	601	BI074785 Ipl_15-B03.B1.A002 Imm	gb_gss:CMS03A3W	8.00	119.86	2.0e+03	959	AL303456 Tetradodon nigrovir
gb_est2:BI074770	8.00	123.64	1.2e+03	601	AV933728 AV933728 K. Sato unpu	gb_gss:CMS03A3W	8.00	119.79	2.0e+03	967	AL251576 Tetradodon nigrovir
gb_gss:AO1156267	8.00	123.63	1.2e+03	602	BI074770 Ipl_15-C06.B1.A002 Imm	gb_gss:AO103651	8.00	119.43	2.1e+03	1011	AG103651 Pan troglodytes DN
gb_gss:AO293578	8.00	123.57	1.2e+03	606	AO1156267 RPCI11-42M4.TK.1 RPCI-	gb_est2:BF4527346	8.00	119.30	2.1e+03	1027	BF4527346 60210117B1 NCI_CG
gb_est2:BI421028	8.00	123.56	1.2e+03	607	AZ993578 2M027BK10R Mouse 10kb	gb_est2:BF4527346	8.00	119.22	2.2e+03	1038	BE477107 PSR6421 ITFC PSR W
gb_gss:AO267250	8.00	123.52	1.2e+03	607	BI421028 LJNES6413R Lotus japc	gb_gss:AO267250	8.00	119.18	2.2e+03	1044	BI433861 Tetradodon nigrovir
gb_gss:AO267250	8.00	123.52	1.2e+03	610	AO267250 CTRB1-E1-2512K4.TR CIT	gb_gss:AO267250	8.00	119.18	2.2e+03	1081	BI433861 Tetradodon nigrovir
gb_est1:AM076591	8.00	123.48	1.2e+03	613	AO267250 nxbx003161F CUCI Rice	gb_est2:BI413348	8.00	118.89	2.4e+03	1149	BI413348 602991833F1 NCI_CG
gb_est1:AM076591	8.00	123.48	1.2e+03	613	AM076591 f133a01.x1 zebrafish a	gb_est2:BI59716	8.00	118.33	2.4e+03	1159	BI59716 603061952F1 NIH_MG
gb_est1:AM052978	8.00	123.48	1.2e+03	613	AM052978 RCI-CR0249-120100-022	gb_gss:AO673471	8.00	118.92	3.3e+03	1159	AG036471 Pan troglodytes DN
gb_gss:BH339082	8.00	123.46	1.2e+03	617	AM052978 RCI-CR0249-290100-022	gb_gss:AO673471	8.00	118.91	3.3e+03	1161	AG071345 Pan troglodytes DN
gb_gss:AO253911	8.00	123.43	1.2e+03	617	BH339082 CH230-105P11.TJ CHORI-	gb_gss:AO253911	8.00	118.42	3.7e+03	1745	AG041888 Pan troglodytes DN
gb_est2:BF291294	8.00	123.39	1.3e+03	620	AZ583911 MHED38A11F Mouse 10kb	gb_hcc:AC0019121	8.00	118.42	4.0e+03	1877	BC019121 Mus musculus, clon
gb_gss:AO2090882	8.00	123.34	1.3e+03	624	BF291294 WHE2221.E05.1092S Aeg	gb_est1:AA771007	8.00	122.62	1.4e+03	81	AA771007 vtl4f10.r1 Barstead
gb_est2:BM684233	8.00	123.25	1.3e+03	625	AZ090882 RPCI-23-458C16.TJ RPCI	gb_est1:AA669656	8.00	122.62	1.4e+03	81	AA669656 1020.Lobliol1F1 pnc C
gb_gss:AO627618	8.00	123.25	1.3e+03	625	BE784253 601471490F1 NIH_MGC-67	gb_est1:AA669656	8.00	118.78	2.0e+03	115	AA669656 1838005.S1 Soares-P
gb_est1:AM0539537	8.00	123.12	1.3e+03	641	AO627618 CTRB1-E1-2651G15.TRB C	gb_gss:AO941132	8.00	118.78	2.0e+03	115	AO941132 Sheared DNA-3564.RP
gb_est1:AM5822118	8.00	123.11	1.3e+03	642	AM0539537 b184b05.w1 B184shear	gb_gss:AO941132	8.00	118.78	2.0e+03	115	AZ715991 RPCI-24-7106.TV RPC
gb_gss:AG1588211	8.00	123.09	1.3e+03	643	AV882118 AV882118 Nori Satoh un	gb_est2:BM441552	8.00	118.50	2.1e+03	119	BM441552 Ema05-S0004.M17-R
gb_est1:AM191806	8.00	123.07	1.3e+03	645	AG155287 Pan troglodytes DNM, c	gb_gss:AG2496903	8.00	118.23	2.1e+03	123	AZ496903 IM0331010R Mouse 10
gb_est2:BM624614	8.00	123.06	1.3e+03	646	AM191806 EST002.CDNAs of Nicot	gb_est2:BI008806	8.00	119.10	2.2e+03	125	BI008806 OV3-RF0066-120201-5
gb_gss:AO3473745	8.00	123.03	1.3e+03	648	BM624614 602728548F1 NIH_MGC-15	gb_est1:BE186927	8.00	119.04	2.2e+03	126	BE186927 946014G04.X1 946-
gb_gss:AO3473745	8.00	123.03	1.3e+03	648	AZ403745 IM0171010R Mouse 10kb	gb_gss:AO294265	8.00	118.79	2.3e+03	130	AZ794265 2M0047P22R Mouse 10
gb_gss:AO347388	8.00	122.99	1.3e+03	651	AZ342810 IM0076C24F Mouse 10kb	gb_est1:AF094600	8.00	118.64	2.3e+03	133	AF094600 AF094600 Mus muscul
gb_est2:BF497118	8.00	122.88	1.3e+03	660	AO347388 RPCI-11-31915.TJ RPCI	gb_gss:AO347388	8.00	118.54	2.3e+03	134	AZ347388 IM0028107R Mouse 10
gb_gss:AO347388	8.00	122.88	1.3e+03	660	BF497118 AT11192.SPrime AT Dros	gb_est2:BF155405	8.00	118.48	2.4e+03	135	BF155405 OV4-HT0538-220900-4
gb_gss:AO347388	8.00	122.88	1.3e+03	660	AZ303313 85.Seg1 Human chromos	gb_gss:AO347388	8.00	118.48	2.4e+03	135	AZ347388 RPCI-24-69H20.TJ RP
gb_gss:AO347388	8.00	122.88	1.3e+03	660	AO347388 RPCI11-101G16.TJB RPCI	gb_est2:DI20273	8.00	118.24	2.4e+03	139	DI20273 HDMS15401.Layer HepG
gb_gss:AO347388	8.00	122.77	1.4e+03	669	BM357742 CH230-199D14.TV CHORI-	gb_est1:AA673288	8.00	117.90	2.5e+03	145	AA673288 v136005.S1 Stratage
gb_est1:AV877330	8.00	122.58	1.4e+03	685	AV877360 AV877360 Nori Satoh un	gb_gss:BM366678	8.00	117.90	2.5e+03	145	BM366678 CH230-2205.TJ CHORI
gb_est1:AV877330	8.00	122.58	1.4e+03	685	BM442449 Eban01-S0003.D11.R IGH	gb_est1:AV106912	8.00	117.63	2.6e+03	150	AV106912 AV106912 Mus muscul
gb_est1:AV877330	8.00	122.44	1.4e+03	697	AV890722 AV890722 Nori Satoh un	gb_gss:AG265750	8.00	117.63	2.6e+03	150	AG265750 IM0533B1R Mouse 10
gb_est2:BI090511	8.00	122.34	1.4e+03	706	BI090511 602855611F1 NIH_MGC-10	gb_est1:AV128660	8.00	117.47	2.7e+03	153	AV128660 OV1-UT0094-121000-4
gb_est2:BI090511	8.00	122.34	1.4e+03	706	BI090511 602855611F1 NIH_MGC-10	gb_est2:BF908914	8.00	117.36	2.7e+03	155	BF908914 RC-BRT122-180399-068
gb_gss:AO2958369	8.00	122.30	1.4e+03	709	AZ958369 2M0225F07R Mouse 10kb	gb_est1:AV017089	8.00	117.31	2.7e+03	156	AV017089 TCBAPAD485F1 Pediatr
gb_gss:AO2958369	8.00	122.29	1.5e+03	710	AG079855 Pan troglodytes DNM, c	gb_est2:BM153510	8.00	117.31	2.7e+03	156	BM153510 TCBAPAD485F1 Pediatr
gb_gss:AO2958369	8.00	122.29	1.5e+03	710	BM380828 AC-ND-174B1.TF ND-TAM	gb_gss:AO2958369	8.00	117.21	2.8e+03	158	AO2958369 p0f11.pK004.F14 no
gb_gss:AO2958369	8.00	122.12	1.5e+03	725	BH471552 CH230-344M.TJ CHORI-23	gb_gss:AO2958369	8.00	117.21	2.8e+03	158	AO2958369 p0f11.pK004.F14 no
gb_gss:AO2958369	8.00	122.10	1.5e+03	727	BH471552 CH230-344M.TJ CHORI-23	gb_gss:AO2958369	8.00	117.21	2.8e+03	158	AO2958369 p0f11.pK004.F14 no
gb_gss:AO2958369	8.00	122.09	1.5e+03	728	AZ186153 SP-1006.A2.G07.SP6 Stc	gb_est1:BB308150	8.00	117.15	2.8e+03	159	BB308150 BB308150 RIKEN full
gb_gss:AO2958369	8.00	122.07	1.5e+03	730	BH487675 BOHEKRT1F BOHE Brasid	gb_est1:AA237232	8.00	117.10	2.8e+03	160	AA237232 MM966d12.L1 Soares m
gb_gss:AO2958369	8.00	122.06	1.5e+03	731	BF261058 HYSME0023K10F Hordeum	gb_gss:AO2958369	8.00	117.05	2.8e+03	161	BH295935 CH230-187B22.TJ CHO
gb_gss:AO2958369	8.00	122.01	1.5e+03	735	BM351398 CH230-175A4.TJ CHORI-2	gb_est2:BI791934	8.00	116.90	2.9e+03	164	BI791934 1610004.S1 HR85 1s1
gb_gss:AO2958369	8.00	121.93	1.5e+03	742	AO687189 nxbx0073M12F CUCI Rice	gb_est1:AA663964	8.00	116.76	2.9e+03	167	AA663964 ae70004.S1 Stratage
gb_gss:AO2958369	8.00	121.91	1.5e+03	744	AG146106 Pan troglodytes DNM, c	gb_gss:BM293471	8.00	116.76	2.9e+03	167	BM293471 CH230-115H15.TVB CH
gb_est2:BF861524	8.00	121.77	1.6e+03	757	BF861524 963023H10.Y1 C. reinha	gb_est2:BE293471	8.00	116.57	3.0e+03	171	BE293471 600942804F1 NIH_MGC
gb_est2:BF215436	8.00	121.76	1.6e+03	758	BF215436 601880966F1 NIH_MGC-57	gb_gss:AO2888194	8.00	116.47	3.1e+03	173	AZ888194 RPCI-24-168A10.TJ R

gb_gss:BH081365	+	7.00	116.38	3.1e+03	175	BH081365	RPCI-24-292N3.TJ	RPCI-24-292N3.TJ	gb_est2:B1040696	+	7.00	114.41	4.0e+03	223	B1040696	CM4-NT0287-120201-8
gb_est2:BF659767	-	7.00	116.33	3.1e+03	176	BF659767	ma31c07.y1	NCI CGAP	gb_est1:AV280455	+	7.00	114.34	4.0e+03	225	AV280455	AV280455
gb_gss:AHV55670	-	7.00	116.10	3.2e+03	179	BH266086	CH230-136D1.TV	CHORI	gb_est1:AV289791	-	7.00	114.34	4.0e+03	225	AV289791	AV289791
gb_est1:AM224371	+	7.00	116.10	3.2e+03	181	AV555670	AV555670	Arabiopsis	gb_est1:AM280832	-	7.00	114.31	4.0e+03	226	AM280832	m180c11.r1
gb_est1:AM280344	+	7.00	116.06	3.2e+03	182	AM724371	uo28c04.x1	NCI CGAP	gb_est1:AM780667	-	7.00	114.31	4.0e+03	226	AM780667	sb15c01.y1
gb_est2:BG82311	+	7.00	115.02	3.2e+03	183	AM730344	vq149c6.r1	Bartsch	gb_est1:BB201443	+	7.00	114.31	4.0e+03	226	BB201443	BB201443
gb_est1:AV039349	+	7.00	115.97	3.3e+03	184	BG82311	PM1-CN0098-150201-011-	mus musculus	gb_est1:BB521544	+	7.00	114.31	4.0e+03	226	BB521544	BB521544
gb_est1:BB046590	-	7.00	115.97	3.3e+03	184	AV039349	AV039349	mus musculus	gb_est1:BB519592	-	7.00	114.24	4.1e+03	228	BB519592	BB519592
gb_est1:AV246346	+	7.00	115.93	3.3e+03	185	BB046590	BB046590	Riken full-16	gb_gss:AA359348	+	7.00	114.24	4.1e+03	228	AA359348	IM0102809F
gb_est1:AM874145	+	7.00	115.88	3.3e+03	186	AV246346	AV246346	Riken full-16	gb_gss:AA359348	+	7.00	114.24	4.1e+03	228	AA359348	IM0102809F
gb_est2:BI189220	+	7.00	115.88	3.3e+03	186	AM874145	hg5c08.x1	NCI CGAP	gb_est1:BB150835	+	7.00	114.16	4.1e+03	230	BB150835	RC3-S70281-121299-0
gb_est1:BI189220	+	7.00	115.88	3.3e+03	186	BI189220	e2e065.r1	Fusarium	gb_est1:AM331151	+	7.00	114.16	4.1e+03	230	AM331151	RC3-S70281-121299-0
gb_est1:BI060066	-	7.00	115.71	3.4e+03	190	BI189220	tc21a05.r1	Soares	gb_est1:BI461091	+	7.00	114.09	4.2e+03	232	BI461091	UI-R-B50-9mwf-01-0
gb_est1:AV092156	-	7.00	115.75	3.4e+03	189	BI060066	IL3-UW0114-16101-419-	mus musculus	gb_est1:AM534811	+	7.00	114.02	4.2e+03	234	AM534811	UI-R-B50-9mwf-01-0
gb_est2:BI060066	-	7.00	115.67	3.4e+03	191	AV092156	AV092156	mus musculus	gb_est1:BB306526	+	7.00	114.02	4.2e+03	234	BB306526	BB306526
gb_gss:BH36347	-	7.00	115.67	3.4e+03	191	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:AM903803	+	7.00	115.63	3.4e+03	193	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:AM903803	+	7.00	115.59	3.4e+03	193	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BF96985	+	7.00	115.54	3.4e+03	194	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BF96985	+	7.00	115.46	3.5e+03	196	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_gss:BH078864	-	7.00	115.30	3.6e+03	200	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:AV284252	+	7.00	115.26	3.6e+03	201	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:AM280521	+	7.00	115.26	3.6e+03	201	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_gss:AM288620	-	7.00	115.22	3.6e+03	202	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_gss:AM288620	-	7.00	115.22	3.6e+03	202	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:BB173809	+	7.00	115.18	3.6e+03	203	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:AM351870	+	7.00	115.14	3.6e+03	204	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_gss:BH507645	-	7.00	115.10	3.6e+03	205	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BG558458	+	7.00	115.06	3.7e+03	206	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_gss:BM232724	-	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est1:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200101-002-	mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	v53605.r1
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BI054282	PM1-GN0210-200									

gb_gss:A2241407	+	7.00	113.46	4.5e+03	251	AZ241407	RPCT-23-89H4.TJ	RPCT-2
gb_gss:BH331295	-	7.00	113.46	4.5e+03	251	BH331295	CH230-53E18.TJ	CHORI-2
gb_est1:A2725878	+	7.00	113.42	4.5e+03	252	A2725878	a125e01.s1	Soares_test
gb_est1:AV275013	-	7.00	113.42	4.5e+03	252	AV275013	AV275013	Riken_full-16
gb_est1:BE183175	-	7.00	113.36	4.5e+03	253	BE183175	RC4-H10664-250400-012-	
gb_est1:AW199563	-	7.00	113.36	4.6e+03	254	AW199563	PM2-DW0053-030300-003-	
gb_est2:B1397796	+	7.00	113.36	4.6e+03	254	B1397796	NKPV_107_F05.F_NXPV (N	
gb_est2:H74827	-	7.00	113.36	4.6e+03	254	H74827	81 Deletion-treated Bras	
gb_est2:F27025	-	7.00	113.33	4.6e+03	255	F27025	HSPD14762_HM3_Homo_Sp16	
gb_gss:BH122334	+	7.00	113.33	4.6e+03	255	BH122334	RPCT-24-273B23.TJ	RPCT
gb_est1:AV055477	-	7.00	113.30	4.6e+03	256	AV055477	AV054777	Mus musculus
gb_est1:AV125212	-	7.00	113.30	4.6e+03	256	AV125212	AV125212	Mus musculus
gb_est2:BM010299	+	7.00	113.30	4.6e+03	256	BM010299	603631178F1_NIH_MGC_41	
gb_est1:AV077072	-	7.00	113.23	4.6e+03	258	AV077072	AV077072	Mus musculus
gb_est2:BE1681709	-	7.00	113.23	4.6e+03	258	BE1681709	461166 MARC_1BOV Bos t	
gb_est1:AV055540	-	7.00	113.20	4.7e+03	259	AV055540	AV055540	Mus musculus
gb_est1:AV129179	-	7.00	113.20	4.7e+03	259	AV129179	AV129179	Mus musculus
gb_est2:F14062	-	7.00	113.20	4.7e+03	259	F14062	ATTS4984_Versailles-VB A	
gb_est1:AA207235	-	7.00	113.17	4.7e+03	260	AA207235	zg73g03.r1	Stratagene
gb_est1:AV128125	-	7.00	113.17	4.7e+03	260	AV128125	AV128125	Mus musculus
gb_est1:AV292569	-	7.00	113.17	4.7e+03	260	AV292569	AV292569	Riken_full-16
gb_est1:BM427331	-	7.00	113.17	4.7e+03	260	BM427331	BM427331	Riken_full-16
gb_est2:BG948366	+	7.00	113.17	4.7e+03	260	BG948366	PM2-KR0007-120101-002-	
gb_est1:BG938534	+	7.00	113.14	4.7e+03	261	BG938534	AV038534	Mus musculus
gb_est2:BG924446	+	7.00	113.14	4.7e+03	261	BG924446	MR3-H10999-020201-002-	
gb_est1:AV313730	-	7.00	113.14	4.7e+03	261	AV313730	AV313730	Mus musculus
gb_est1:AV313730	-	7.00	113.14	4.7e+03	261	AV313730	AV313730	Mus musculus
gb_est1:AV293845	+	7.00	113.11	4.7e+03	262	AV293845	AV293845	Riken_full-16
gb_est1:AA407712	-	7.00	113.11	4.7e+03	262	AA407712	ESTN0966	Mouse 7.5 dpc
gb_gss:BA286E10Q	-	7.00	113.11	4.7e+03	262	BA286E10Q	T. brucei	sheared gend
gb_est1:AV016067	-	7.00	113.08	4.7e+03	263	AV016067	AV016067	Mus musculus
gb_est1:AV120233	-	7.00	113.08	4.7e+03	263	AV120233	AV120233	Mus musculus
gb_est1:BB279453	-	7.00	113.08	4.7e+03	263	BB279453	BB279453	Riken_full-16
gb_est1:BB384755	-	7.00	113.08	4.7e+03	263	BB384755	BB384755	Riken_full-16
gb_est2:BM098347	-	7.00	113.08	4.7e+03	263	BM098347	EBem08_S0002.J13.R	IGH
gb_est2:BM098347	-	7.00	113.08	4.7e+03	263	BM098347	UT-R-CA0-bhy-h-04-0-U1	
gb_est1:AV240748	+	7.00	113.05	4.7e+03	264	AV240748	AV240748	Riken_full-16
gb_est1:AA260072	+	7.00	113.05	4.7e+03	264	AA260072	va38g11.r1	Soares_mous
gb_est1:BB369966	-	7.00	113.05	4.7e+03	264	BB369966	BB369966	Riken_full-16
gb_est1:FL7024	-	7.00	113.05	4.7e+03	264	FL7024	HSPD01539_Hall_Homo_Sp16	
gb_gss:BA2912545	-	7.00	113.05	4.7e+03	264	BA2912545	RPCT-24-164P14.TJ	RGPI
gb_est1:AA108778	-	7.00	113.02	4.8e+03	265	AA108778	m35b07.r1	Barstead_MF
gb_est2:BF847555	+	7.00	113.02	4.8e+03	265	BF847555	CM3-ENM008-311000-441-	
gb_est1:AV313255	-	7.00	112.99	4.8e+03	266	AV313255	AV313255	Riken_full-16
gb_est1:AV362557	-	7.00	112.99	4.8e+03	266	AV362557	AV362527	Riken_full-16
gb_est2:BF290155	-	7.00	112.99	4.8e+03	266	BF290155	EST9454746	Rat Gene Ind
gb_est1:AV042913	-	7.00	112.93	4.8e+03	268	AV042913	AV042913	Mus musculus
gb_est2:BE925248	+	7.00	112.93	4.8e+03	268	BE925248	CM4-ANM081-210800-274-	
gb_est2:H32104	+	7.00	112.90	4.8e+03	269	H32104	EST106901	Rat PC-12 cell
gb_gss:AA077876	+	7.00	112.90	4.8e+03	269	AA077876	CIT-HSP-236662.TF	CIT
gb_est1:AA148809	+	7.00	112.87	4.9e+03	270	AA148809	mK09a06.x1	Soares_mous
gb_est1:AA1659816	-	7.00	112.87	4.9e+03	270	AA1659816	AV03C09.x1	NCI_CGAP_Pt
gb_est1:AV127766	-	7.00	112.87	4.9e+03	270	AV127766	AV127766	Mus musculus
gb_gss:BH33406	-	7.00	112.87	4.9e+03	270	BH33406	HS-1017-B1-D04-MF	abi CT
gb_gss:BH333073	-	7.00	112.87	4.9e+03	270	BH333073	CH230-202A18.TJ	CHORI
gb_est2:BE477362	-	7.00	112.84	4.9e+03	271	BE477362	us1d08.y1	Perkins_LRH
gb_est1:AA303636	-	7.00	112.81	4.9e+03	272	AA303636	EST16609	Aortic endothel
gb_est1:AW945614	-	7.00	112.81	4.9e+03	272	AW945614	RC2-ENM027-290400-011-	
gb_gss:AA46929	-	7.00	112.75	4.9e+03	273	AA46929	IM0243M02R	Mouse 10kb
gb_est1:AA330073	-	7.00	112.75	4.9e+03	274	AA330073	EST22608	Adipose_tissu
gb_est1:AA320146	+	7.00	112.75	4.9e+03	274	AA320146	EST22534	Adipose_tissu
gb_est1:BB418055	+	7.00	112.75	4.9e+03	274	BB418055	BB418055	Riken_full-16
gb_gss:AA255056	+	7.00	112.75	4.9e+03	274	AA255056	PC2-209C13.TV	RPCT
gb_gss:AA099522	+	7.00	112.72	4.9e+03	275	AA099522	HS_3049.42	E08_Rat CIT
gb_est1:BB161070	+	7.00	112.69	5.0e+03	276	BB161070	BB161070	Riken_full-16
gb_est1:AA501687	+	7.00	112.66	5.0e+03	277	AA501687	ng06e08.s1	NCI_CGAP_L1
gb_est2:BI072442	+	7.00	112.66	5.0e+03	277	BI072442	C076P08U	Populus_stra1
gb_gss:BH383224	+	7.00	112.63	5.0e+03	278	BH383224	AG-ND-150P19.TR	ND-TAM
gb_gss:BP97100	+	7.00	112.60	5.0e+03	279	BP97100	T31M10TF	TAMU_Arden
gb_est1:AA1318342	+	7.00	112.60	5.0e+03	279	AA1318342	ta73c09.x1	NCI_CGAP_H9
gb_est1:AA332974	+	7.00	112.60	5.0e+03	279	AA332974	AV332974	Riken_full-16
gb_est1:BB010325	-	7.00	112.60	5.0e+03	279	BB010325	BB010325	Riken_full-16
gb_est1:BB353901	-	7.00	112.60	5.0e+03	279	BB353901	BB353901	Riken_full
gb_est2:BF33135	+	7.00	112.60	5.0e+03	279	BF33135	EST108835	Rat PC-12 c
gb_est2:BF331351	-	7.00	112.60	5.0e+03	279	BF331351	OV73-H10457-060400-01	
gb_est1:AV074256	-	7.00	112.57	5.0e+03	280	AV074256	AV074256	Mus muscul
gb_est1:BE004191	-	7.00	112.57	5.0e+03	280	BE004191	CM0-BM0103-180300-2	
gb_est1:AV057774	-	7.00	112.54	5.1e+03	281	AV057774	AV057774	Mus muscul
gb_est1:BB434311	-	7.00	112.54	5.1e+03	281	BB434311	BB434311	Riken_full
gb_est2:BM435255	-	7.00	112.54	5.1e+03	281	BM435255	16u1383.abi	Bos tau
gb_est1:AA036547	-	7.00	112.51	5.1e+03	282	AA036547	AV036547	Right hemi
gb_est1:AA036846	-	7.00	112.51	5.1e+03	282	AA036846	EST556567	Right hemi
gb_est1:BB230078	-	7.00	112.51	5.1e+03	282	BB230078	BB230078	Riken_full
gb_gss:BP98042	-	7.00	112.51	5.1e+03	282	BP98042	F23M1878B	IGF Arabidid
gb_est1:AM308582	-	7.00	112.49	5.1e+03	283	AM308582	SNVC3MCA36G01SK	On
gb_gss:BB166916	-	7.00	112.49	5.1e+03	283	BB166916	SLAK_000344	Arabidid
gb_est1:AA355960	-	7.00	112.46	5.1e+03	284	AA355960	EST64438	Jurkat T-C
gb_est1:BB154575	-	7.00	112.46	5.1e+03	284	BB154575	BB154575	Riken_full
gb_est2:BF889065	+	7.00	112.46	5.1e+03	284	BF889065	RC4-TM0130-281100-0	
gb_est1:AV035655	-	7.00	112.43	5.1e+03	285	AV035655	AV035655	Mus muscul
gb_est1:AV209994	-	7.00	112.43	5.1e+03	285	AV209994	AV209994	Riken_full
gb_est1:BB003919	-	7.00	112.43	5.1e+03	285	BB003919	BB003919	Riken_full
gb_est1:BB009866	-	7.00	112.43	5.1e+03	285	BB009866	BB009866	Riken_full
gb_est1:BB195185	-	7.00	112.43	5.1e+03	285	BB195185	BB195185	Riken_full
gb_est1:BB287796	-	7.00	112.37	5.2e+03	287	BB287796	BB287796	Riken_full
gb_est1:AV106104	-	7.00	112.37	5.2e+03	287	AV106104	AV106104	Mus muscul
gb_est1:BB233066	-	7.00	112.37	5.2e+03	287	BB233066	BB233066	Riken_full
gb_est1:AA619868	-	7.00	112.37	5.2e+03	287	AA619868	V158F07.s1	Knowles
gb_est2:BI069236	-	7.00	112.37	5.2e+03	287	BI069236	C034P21U	Populus st
gb_est2:BM115451	-	7.00	112.37	5.2e+03	287	BM115451	L0820C01-3	NIA Mous
gb_est2:BF469294	-	7.00	112.37	5.2e+03	287	BF469294	UT-M-BH3-aug-d-12-0	
gb_est1:AV037189	-	7.00	112.34	5.2e+03	288	AV037189	AV037189	Mus muscul
gb_est1:BB292094	-	7.00	112.34	5.2e+03	288	BB292094	BB292094	Riken_full
gb_est1:BB301775	-	7.00	112.34	5.2e+03	288	BB301775	BB301775	Riken_full
gb_est1:AV125360	-	7.00	112.32	5.2e+03	289	AV125360	AV125360	Mus muscul
gb_est1:AV129353	-	7.00	112.32	5.2e+03	289	AV129353	AV129353	Mus muscul
gb_est1:AV149430	-	7.00	112.32	5.2e+03	289	AV149430	AV149430	Mus muscul
gb_est1:AV156270	-	7.00	112.32	5.2e+03	289	AV156270	AV156270	Mus muscul
gb_est1:BB349031	-	7.00	112.32	5.2e+03	289	BB349031	BB349031	Riken_full
gb_est1:AV000779	-	7.00	112.29	5.2e+03	290	AV000779	AV000779	Mus muscul
gb_est1:AV149267	-	7.00	112.29	5.2e+03	290	AV149267	AV149267	Mus muscul
gb_est1:AM308378	-	7.00	112.29	5.2e+03	290	AM308378	3316 MARC_1BOV Bos	
gb_est1:BB006807	-	7.00	112.29	5.2e+03	290	BB006807	BB006807	Riken_full
gb_est1:BB232108	-	7.00	112.29	5.2e+03	290	BB232108	BB232108	Riken_full
gb_est2:D62781	-	7.00	112.29	5.2e+03	290	D62781	HUM245C474	Human Embryo
gb_gss:BB314813	-	7.00	112.29	5.2e+03	290	BB314813	CH230-40M14.TJ	CHOR
gb_est1:AV037202	-	7.00	112.26	5.3e+03	291	AV037202	AV037202	Riken_full
gb_est1:AV299926	-	7.00	112.26	5.3e+03	291	AV299926	AV299926	Riken_full
gb_est1:BB388463	-	7.00	112.26	5.3e+03	291	BB388463	BB388463	Riken_full
gb_gss:BB355521	-	7.00	112.26	5.3e+03	291	BB355521	BB604588	Riken_full
gb_est1:AA089106	-	7.00	112.26	5.3e+03	292	AA089106	AV089106	Mus muscul
gb_est1:AA133442	-	7.00	112.23	5.3e+03	292	AA133442	AV133442	Mus muscul
gb_est1:BB478372	-	7.00	112.23	5.3e+03	292	BB478372	BB478372	Riken_full
gb_est1:AA208193	-	7.00	112.20	5.3e+03	293	AA208193	AA208193	Riken_full

gb_gss: A2364844	7.00	112.12	5.3e+03	296	A2364844 IM011113f Mouse 10kb	gb_est2: C35334	7.00	112.01	5.4e+03	300	C35334 C35334 Yuj1 Kohara un
gb_est1: AV034688	7.00	112.09	5.4e+03	297	AV034688 AV034688 Mus musculus	gb_est2: C35340	7.00	112.01	5.4e+03	300	C35340 C35340 Yuj1 Kohara un
gb_est1: BB339322	7.00	112.09	5.4e+03	297	BB339322 BB339322 Riken full-1f	gb_est2: C36101	7.00	112.01	5.4e+03	300	C36101 C36101 Yuj1 Kohara un
gb_est1: BB555278	7.00	112.09	5.4e+03	297	BB555278 BB555278 Riken full-1f	gb_est2: C36780	7.00	112.01	5.4e+03	300	C36780 C36780 Yuj1 Kohara un
gb_est2: T01098	7.00	112.09	5.4e+03	297	T01098 WEST01819 Early embryo,	gb_est2: C36806	7.00	112.01	5.4e+03	300	C36806 C36806 Yuj1 Kohara un
gb_est1: AV104510	7.00	112.07	5.4e+03	298	AV104510 AV104510 Mus musculus	gb_est2: C36846	7.00	112.01	5.4e+03	300	C36846 C36846 Yuj1 Kohara un
gb_est1: BV148368	7.00	112.07	5.4e+03	298	AV148368 AV148368 Mus musculus	gb_est2: C36957	7.00	112.01	5.4e+03	300	C36957 C36957 Yuj1 Kohara un
gb_est1: BV171213	7.00	112.07	5.4e+03	298	BB171213 BB171213 Riken full-1f	gb_est2: C37479	7.00	112.01	5.4e+03	300	C37479 C37479 Yuj1 Kohara un
gb_est2: BF651535	7.00	112.07	5.4e+03	298	BF651535 dd82f01. y1 Wellcome CR	gb_est2: C37774	7.00	112.01	5.4e+03	300	C37774 C37774 Yuj1 Kohara un
gb_est1: AV040324	7.00	112.04	5.4e+03	299	AV040324 AV040324 Mus musculus	gb_est2: C38064	7.00	112.01	5.4e+03	300	C38064 C38064 Yuj1 Kohara un
gb_est1: BB074970	7.00	112.04	5.4e+03	299	BB074970 BB074970 Riken full-1f	gb_est2: C38333	7.00	112.01	5.4e+03	300	C38333 C38333 Yuj1 Kohara un
gb_est1: BB217267	7.00	112.04	5.4e+03	299	BB217267 BB217267 Riken full-1f	gb_est2: C38330	7.00	112.01	5.4e+03	300	C38330 C38330 Yuj1 Kohara un
gb_est1: BB364408	7.00	112.04	5.4e+03	299	BB364408 BB364408 Riken full-1f	gb_est2: C38433	7.00	112.01	5.4e+03	300	C38433 C38433 Yuj1 Kohara un
gb_est1: BB554257	7.00	112.04	5.4e+03	299	BB554257 BB554257 Riken full-1f	gb_est2: C39151	7.00	112.01	5.4e+03	300	C39151 C39151 Yuj1 Kohara un
gb_est1: AV098764	7.00	112.01	5.4e+03	300	AV098764 AV098764 Sugeno Homo S	gb_gss: A244428	7.00	112.01	5.4e+03	300	A244428 IM0239p10r Mouse 10
gb_est1: AV113694	7.00	112.01	5.4e+03	300	AV113694 AV113694 unpublished c	gb_est2: H33584	7.00	111.99	5.4e+03	301	H33584 EST109725 Rat PC-12 c
gb_est1: AV115487	7.00	112.01	5.4e+03	300	AV115487 AV115487 unpublished c	gb_est1: BB207599	7.00	111.96	5.5e+03	302	BB207599 BB207599 Riken full1
gb_est1: AV096384	7.00	112.01	5.4e+03	300	AV096384 AV096384 Mus musculus	gb_est1: BB223027	7.00	111.96	5.5e+03	302	BB223027 BB223027 Riken full1
gb_est1: AV151945	7.00	112.01	5.4e+03	300	AV151945 AV151945 Mus musculus	gb_est1: BB516815	7.00	111.96	5.5e+03	302	BB516815 BB516815 Riken full1
gb_est1: AV175870	7.00	112.01	5.4e+03	300	AV175870 AV175870 Yuj1 Kohara u	gb_est1: AA221917	7.00	111.93	5.5e+03	303	AA221917 mw16h01. r1 Soares m
gb_est1: AV176191	7.00	112.01	5.4e+03	300	AV176191 AV176191 Yuj1 Kohara u	gb_est1: BB120926	7.00	111.93	5.5e+03	303	BB120926 BB120926 Riken full1
gb_est1: AV176665	7.00	112.01	5.4e+03	300	AV176665 AV176665 Yuj1 Kohara u	gb_est1: BB477842	7.00	111.93	5.5e+03	303	BB477842 BB477842 Riken full1
gb_est1: AV176836	7.00	112.01	5.4e+03	300	AV176836 AV176836 Yuj1 Kohara u	gb_est1: BB809793	7.00	111.93	5.5e+03	303	BB809793 BB809793 Riken full1
gb_est1: AV176962	7.00	112.01	5.4e+03	300	AV176962 AV176962 Yuj1 Kohara u	gb_gss: A2465428	7.00	111.93	5.5e+03	303	A2465428 IM0275b18f Mouse 10
gb_est1: AV177696	7.00	112.01	5.4e+03	300	AV177696 AV177696 Yuj1 Kohara u	gb_est1: AV055296	7.00	111.91	5.5e+03	304	AV055296 AV055296 Mus muscul
gb_est1: AV178568	7.00	112.01	5.4e+03	300	AV178568 AV178568 Yuj1 Kohara u	gb_est1: AV082633	7.00	111.91	5.5e+03	304	AV082633 AV082633 Mus muscul
gb_est1: AV179041	7.00	112.01	5.4e+03	300	AV179041 AV179041 Yuj1 Kohara u	gb_est1: AV121146	7.00	111.91	5.5e+03	304	AV121146 AV121146 Mus muscul
gb_est1: AV179289	7.00	112.01	5.4e+03	300	AV179289 AV179289 Yuj1 Kohara u	gb_est2: D62466	7.00	111.91	5.5e+03	304	D62466 HMC290D07B Clontech h
gb_est1: AV179316	7.00	112.01	5.4e+03	300	AV179316 AV179316 Yuj1 Kohara u	gb_est2: Z44674	7.00	111.91	5.5e+03	304	Z44674 HSC265031 normalized
gb_est1: AV179951	7.00	112.01	5.4e+03	300	AV179951 AV179951 Yuj1 Kohara u	gb_gss: A0907967	7.00	111.91	5.5e+03	304	A0907967 GSTR053344 Trypanos
gb_est1: AV180168	7.00	112.01	5.4e+03	300	AV180168 AV180168 Yuj1 Kohara u	gb_est2: BH294987	7.00	111.91	5.5e+03	304	BH294987 CH230-44L16 Tj CHOR
gb_est1: AV180641	7.00	112.01	5.4e+03	300	AV180641 AV180641 Yuj1 Kohara u	gb_est1: BH465143	7.00	111.88	5.5e+03	305	BH465143 mw74h09. x1 Soares m
gb_est1: AV180813	7.00	112.01	5.4e+03	300	AV180813 AV180813 Yuj1 Kohara u	gb_est1: A1561735	7.00	111.88	5.5e+03	305	A1561735 v662h03. x1 Soares t
gb_est1: AV180845	7.00	112.01	5.4e+03	300	AV180845 AV180845 Yuj1 Kohara u	gb_est1: AV060593	7.00	111.88	5.5e+03	305	AV060593 AV060593 Mus muscul
gb_est1: AV180962	7.00	112.01	5.4e+03	300	AV180962 AV180962 Yuj1 Kohara u	gb_est1: AV099888	7.00	111.88	5.5e+03	305	AV099888 AV099888 Mus muscul
gb_est1: AV181070	7.00	112.01	5.4e+03	300	AV181070 AV181070 Yuj1 Kohara u	gb_est2: BE451733	7.00	111.88	5.5e+03	305	BE451733 NXC1_001.C08_F NCBI
gb_est1: AV181082	7.00	112.01	5.4e+03	300	AV181082 AV181082 Yuj1 Kohara u	gb_est2: BE928966	7.00	111.88	5.5e+03	305	BE928966 PM3-GN0012-290800-0
gb_est1: AV181124	7.00	112.01	5.4e+03	300	AV181124 AV181124 Yuj1 Kohara u	gb_est2: BE928999	7.00	111.85	5.5e+03	306	BE928999 PM3-GN0012-290800-0
gb_est1: AV181807	7.00	112.01	5.4e+03	300	AV181807 AV181807 Yuj1 Kohara u	gb_est1: AV038497	7.00	111.85	5.5e+03	306	AV038497 AV038497 Dictyostel
gb_est1: AV182339	7.00	112.01	5.4e+03	300	AV182339 AV182339 Yuj1 Kohara u	gb_est1: AM540739	7.00	111.83	5.6e+03	307	AM540739 C0136P03-3 NIA Mous
gb_est1: AV182749	7.00	112.01	5.4e+03	300	AV182749 AV182749 Yuj1 Kohara u	gb_est1: AV027493	7.00	111.83	5.6e+03	308	AV027493 AV027493 Mus muscul
gb_est1: AV183431	7.00	112.01	5.4e+03	300	AV183431 AV183431 Yuj1 Kohara u	gb_est1: AV166744	7.00	111.80	5.6e+03	308	AV166744 AV166744 Mus muscul
gb_est1: AV183656	7.00	112.01	5.4e+03	300	AV183656 AV183656 Yuj1 Kohara u	gb_gss: BH159127	7.00	111.80	5.6e+03	308	BH159127 CH230-GEF22.TV CHORI
gb_est1: AV183691	7.00	112.01	5.4e+03	300	AV183691 AV183691 Yuj1 Kohara u	gb_gss: TA050805P	7.00	111.80	5.6e+03	308	TA050805P T. brucei sheared g
gb_est1: AV184110	7.00	112.01	5.4e+03	300	AV184110 AV184110 Yuj1 Kohara u	gb_est1: A1268654	7.00	111.77	5.6e+03	309	A1268654 g039g01. x1 NCI_CGAP
gb_est1: AV184698	7.00	112.01	5.4e+03	300	AV184698 AV184698 Yuj1 Kohara u	gb_est1: AM823376	7.00	111.77	5.6e+03	309	AM823376 BM203028 Riken full1
gb_est1: AV184870	7.00	112.01	5.4e+03	300	AV184870 AV184870 Yuj1 Kohara u	gb_est1: BB203028	7.00	111.77	5.6e+03	309	BB203028 BB203028 Riken full1
gb_est1: AV185561	7.00	112.01	5.4e+03	300	AV185561 AV185561 Yuj1 Kohara u	gb_est1: BB607366	7.00	111.75	5.6e+03	309	BB607366 BB607366 Riken full1
gb_est1: AV880761	7.00	112.01	5.4e+03	300	AV880761 AV880761 Nori Satoch un	gb_est1: BB500709	7.00	111.75	5.6e+03	310	BB500709 BB500709 Riken full1
gb_est1: BB241194	7.00	112.01	5.4e+03	300	BB241194 BB241194 Riken full-1f	gb_est1: AV112515	7.00	111.72	5.6e+03	311	AV112515 AV112515 Mus muscul
gb_est1: C07987	7.00	112.01	5.4e+03	300	C07987 C07987 Yuj1 Kohara unpun	gb_est1: BB165241	7.00	111.72	5.6e+03	311	BB165241 PM1-HT0521-050500-0
gb_est2: C07987	7.00	112.01	5.4e+03	300	C29424 C29424 Yuj1 Kohara unpun	gb_est2: BE646285	7.00	111.70	5.6e+03	312	BE646285 7684b06. x1 NCI_CGAP
gb_est2: C296324	7.00	112.01	5.4e+03	300	C29632 C29632 Yuj1 Kohara unpun	gb_gss: A20706709	7.00	111.70	5.6e+03	312	A20706709 RCH2-23-421C16.TV R
gb_est2: C29665	7.00	112.01	5.4e+03	300	C29665 C29665 Yuj1 Kohara unpun	gb_gss: BH35376	7.00	111.59	5.7e+03	316	BH35376 CH230-175624.TV CHO
gb_est2: C29759	7.00	112.01	5.4e+03	300	C29759 C29759 Yuj1 Kohara unpun	gb_est1: BB216732	7.00	111.67	5.7e+03	313	BB216732 T95E7Sz672h09. y1 TGR
gb_est2: C29857	7.00	112.01	5.4e+03	300	C29857 C29857 Yuj1 Kohara unpun	gb_gss: PM009615R	7.00	111.62	5.7e+03	315	PM009615R Parameticon tetraure
gb_est2: C29959	7.00	112.01	5.4e+03	300	C29959 C29959 Yuj1 Kohara unpun	gb_est1: A2708876	7.00	111.54	5.8e+03	318	A2708876 v16f61. r1 Soares m
gb_est2: C300318	7.00	112.01	5.4e+03	300	C300318 C300318 Yuj1 Kohara unpun	gb_gss: BH35376	7.00	111.54	5.8e+03	318	BH35376 CH230-175624.TV CHO
gb_est2: C30410	7.00	112.01	5.4e+03	300	C30410 C30410 Yuj1 Kohara unpun	gb_est1: AA385856	7.00	111.57	5.7e+03	317	AA385856 EST99633 Thyroid Ho
gb_est2: C30517	7.00	112.01	5.4e+03	300	C30517 C30517 Yuj1 Kohara unpun	gb_gss: A1737362	7.00	111.54	5.8e+03	318	A1737362 ZM0001007R Mouse 10
gb_est2: C30651	7.00	112.01	5.4e+03	300	C30651 C30651 Yuj1 Kohara unpun	gb_est1: AV130072	7.00	111.54	5.8e+03	318	AV130072 AV130072 Mus muscul
gb_est2: C30783	7.00	112.01	5.4e+03	300	C30783 C30783 Yuj1 Kohara unpun	gb_est1: AV208145	7.00	111.54	5.8e+03	318	AV208145 AV208145 Mus muscul
gb_est2: C30806	7.00	112.01	5.4e+03	300	C30806 C30806 Yuj1 Kohara unpun	gb_est1: AA288462	7.00	111.54	5.8e+03	318	AA288462 v16f61. r1 Soares m
gb_est2: C31209	7.00	112.01	5.4e+03	300	C31209 C31209 Yuj1 Kohara unpun	gb_gss: BH321994	7.00	111.54	5.8e+03	318	BH321994 CH230-173423.TJ CHO
gb_est2: C32518	7.00	112.01	5.4e+03	300	C32518 C32518 Yuj1 Kohara unpun	gb_est1: A141944	7.00	111.52	5.8e+03	319	A141944 s983a807. y1 Gm-cl004
gb_est2: C32966	7.00	112.01	5.4e+03	300	C32966 C32966 Yuj1 Kohara unpun	gb_est1: A1725172	7.00	111.52	5.8e+03	319	A1725172 1071.P1RfG2.Plus1-t
gb_est2: C32994	7.00	112.01	5.4e+03	300	C32994 C32994 Yuj1 Kohara unpun	gb_est2: BE986015	7.00	111.52	5.8e+03	319	BE986015 CML-DT0048-1701u1-7
gb_est2: C33492	7.00	112.01	5.4e+03	300	C33492 C33492 Yuj1 Kohara unpun	gb_est2: BE944188	7.00	111.52	5.8e+03	319	BE944188 U1-M-B83-wn-b-05-0
gb_est2: C34761	7.00	112.01	5.4e+03	300	C34761 C34761 Yuj1 Kohara unpun	gb_est1: AA064282	7.00	111.49	5.8e+03	320	AA064282 m167f06. r1 Soares m
gb_est2: C34875	7.00	112.01	5.4e+03	300	C34875 C34875 Yuj1 Kohara unpun	gb_est1: A1034580	7.00	111.49	5.8e+03	320	A1034580 Lmk339p10/252c Lets
gb_est2: C35145	7.00	112.01	5.4e+03	300	C35145 C35145 Yuj1 Kohara unpun	gb_est1: BB136940	7.00	111.49	5.8e+03	320	BB136940 BB136940 Riken full1
gb_est2: C35185	7.00	112.01	5.4e+03	300	C35185 C35185 Yuj1 Kohara unpun	gb_est1: BB140572	7.00	111.49	5.8e+03	320	BB140572 BB140572 Riken full1

gb_est2:BF802920	-	7.00	111.49	5.8e+03	320	BF802920	IL5-C10149-011100-222-	gb_est1:AV209743	+	7.00	110.90	6.2e+03	344	AV209743	AV209743	RIKEN	full1
gb_est2:BI133330	+	7.00	111.49	5.8e+03	320	BI133330	UI-M-BH3-brv-g-06-0-01	gb_est2:BF339943	+	7.00	110.90	6.2e+03	344	BF339943	UI-R-CAL-Bjb-h-17-0		
gb_est1:AA736935	+	7.00	111.46	5.8e+03	321	AA736935	oa48c03.s1 NC1 CGAP GC	gb_est1:AV011452	+	7.00	110.86	6.3e+03	346	AV011452	AV011452	Mus	muscul
gb_est1:AAV939711	+	7.00	111.46	5.8e+03	321	AAV939711	476 PIPIFG plus tede	gb_est2:237589	+	7.00	110.86	6.3e+03	346	237589	ATTS4054	Strasbourg-A	
gb_est1:AAV903016	+	7.00	111.46	5.8e+03	321	AAV903016	AV090316	gb_est2:BE748459	+	7.00	110.86	6.3e+03	346	BE748459	601571953P1	NIH-MGC	
gb_est2:DJ13307	+	7.00	111.46	5.8e+03	321	DJ13307	CELMK67CYR	gb_gss:AE514891	+	7.00	110.86	6.3e+03	346	AE514891	IM0361E24R	Mouse	10
gb_est1:AA122513	+	7.00	111.44	5.8e+03	322	AA12513	ma77d04.r1 Soares	gb_est1:BI202100	+	7.00	110.83	6.3e+03	347	BI202100	q132604.r1	NC1 CGAP	
gb_est1:AAW012513	+	7.00	111.44	5.8e+03	322	AAW012513	IL5-DM0067-030400-053-	gb_est2:BE638842	+	7.00	110.83	6.3e+03	347	BE638842	RPCT-23-209113	TV R	
gb_est1:BB1440573	-	7.00	111.44	5.8e+03	322	BB1440573	BB1440573	gb_gss:BA875542	+	7.00	110.83	6.3e+03	347	BA875542	RPCT-23-209113	TV R	
gb_est2:BM1395032	-	7.00	111.44	5.8e+03	322	BM1395032	k165f04.y1 Ascatis suu	gb_gss:BA8779	+	7.00	110.81	6.3e+03	348	BA8779	RPCT-23-209113	TV R	
gb_est2:HB98464	+	7.00	111.44	5.8e+03	322	HB98464	YMO4e05.r1 Soares melanc	gb_est2:BF542284	+	7.00	110.81	6.3e+03	348	BF542284	UI-R-C2-rs-C-06-0-0		
gb_gss:AZ706922	-	7.00	111.44	5.8e+03	322	AZ706922	RPCT-23-23682.TV RCT	gb_est1:BA441046	+	7.00	110.79	6.3e+03	349	BA441046	v166903.r1 Soares	Pla	
gb_gss:BM363863	-	7.00	111.44	5.8e+03	322	BM363863	CH230-1190H2.TV RCT	gb_est2:R23372	+	7.00	110.79	6.3e+03	349	R23372	yh32901.r1 Soares	Pla	
gb_est1:AV161696	-	7.00	111.41	5.9e+03	323	AV161696	AV161696	gb_gss:AZ383245	+	7.00	110.79	6.3e+03	349	AZ383245	IM0149104R	Mouse	10
gb_est1:AA1900133	-	7.00	111.31	5.9e+03	324	AA1900133	ws29e08.x1 NC1 CGAP GC	gb_gss:AO247626	+	7.00	110.79	6.3e+03	349	AO247626	HS-2055.B2.F09.T7 C		
gb_est2:DJ17465	+	7.00	111.36	5.9e+03	324	DJ17465	CELMK075GIR	gb_gss:AO294832	+	7.00	110.79	6.3e+03	349	AO294832	HS-2183.B1.H12.MF C		
gb_est1:BB2424230	+	7.00	111.36	5.9e+03	324	BB2424230	BB2424230	gb_gss:AO297085	+	7.00	110.79	6.3e+03	349	AO297085	HS-3034.B2.C10.MF C		
gb_est2:BG982818	+	7.00	111.34	5.9e+03	326	BG982818	PMO-CM0150-010301-001	gb_est2:BG053004	+	7.00	110.76	6.4e+03	350	BG053004	H3003403-3	NIA Mous	
gb_est2:BE517111	+	7.00	111.34	5.9e+03	326	BE517111	WHE0624.C07.F142A	gb_est1:BA638942	+	7.00	110.74	6.4e+03	351	BA638942	AV528892	Arbidiops	1
gb_gss:AZ617949	+	7.00	111.31	5.9e+03	327	AZ617949	IM044024F	gb_est1:AV528892	+	7.00	110.74	6.4e+03	351	AV528892	AV528892	Arbidiops	1
gb_est1:AA408233	-	7.00	111.31	5.9e+03	327	AA408233	EST02735	gb_est1:BE145088	+	7.00	110.74	6.4e+03	351	BE145088	CM3-RH0192-071059-0		
gb_gss:AO918294	+	7.00	111.31	5.9e+03	327	AO918294	RPCT-23-287R8.TV RCT	gb_est1:BB145088	+	7.00	110.74	6.4e+03	351	BB145088	RPCT-23-324H15	TV R	
gb_gss:BM314949	-	7.00	111.29	5.9e+03	327	BM314949	CH230-119G16.TV CHOR	gb_gss:AO933894	+	7.00	110.74	6.4e+03	351	AO933894	RPCT-23-324H15	TV R	
gb_est1:BB640646	+	7.00	111.29	5.9e+03	328	BB640646	BB640646	gb_gss:AO986534	+	7.00	110.74	6.4e+03	351	AO986534	RPCT-23-324H15	TV R	
gb_gss:BB809585	+	7.00	111.29	5.9e+03	328	BB809585	BB809585	gb_gss:AO301772	+	7.00	110.74	6.4e+03	351	AO301772	HS-3041.B1.A03.LR C		
gb_est2:BI072356	+	7.00	111.29	5.9e+03	328	BI072356	CM4-MW0247-190101-815-	gb_est2:BI338648	+	7.00	110.72	6.4e+03	352	BI338648	MI-P-AY1-npv-c-09-0		
gb_est2:HB87742	+	7.00	111.29	5.9e+03	328	HB87742	y574f06.s1 Soares retina	gb_est1:BI824517	+	7.00	110.70	6.4e+03	353	BI824517	cx17c08.x1 NC1 CGAP		
gb_est2:BE522777	+	7.00	111.27	5.9e+03	328	BE522777	M286588	gb_est1:AU070993	+	7.00	110.70	6.4e+03	353	AU070993	AU070993	Stratagene	
gb_est2:DJ272554	+	7.00	111.27	6.0e+03	329	DJ272554	CELMK104F8R	gb_est2:J94714	+	7.00	110.67	6.4e+03	353	J94714	y635610.r1	Stratagene	
gb_est1:AA106152	+	7.00	111.24	6.0e+03	330	AA106152	hb02a12.p1 ZP adult	gb_est1:MW346576	+	7.00	110.67	6.4e+03	354	MW346576	HS-60711	seq.f Human	
gb_est1:AV547237	+	7.00	111.24	6.0e+03	330	AV547237	AV547237	gb_gss:BM324106	+	7.00	110.67	6.4e+03	354	BM324106	CH230-15624.TV CHOR		
gb_est2:BF839824	+	7.00	111.24	6.0e+03	330	BF839824	QVI-MW0170-131100-459-	gb_est1:AV014209	+	7.00	110.65	6.5e+03	355	AV014209	AV014209	Mus	muscul
gb_gss:AA074727	+	7.00	111.24	6.0e+03	330	AA074727	CIT-HSP-23514A.TF CIT	gb_gss:AM747893	+	7.00	110.65	6.5e+03	355	AM747893	QVI-BM0001-301199-0		
gb_gss:BM171755	+	7.00	111.24	6.0e+03	330	BM171755	SALK-004749	gb_gss:AO650478	+	7.00	110.65	6.5e+03	355	AO650478	Sheared DNA-20J5.TR		
gb_est1:AA145738	+	7.00	111.17	6.0e+03	332	AA145738	ga34b09.y1 Moss EST 13	gb_est1:AA247820	+	7.00	110.63	6.5e+03	356	AA247820	UI-HF0711	seq.f Human	
gb_est1:BB144612	+	7.00	111.17	6.0e+03	332	BB144612	BB144612	gb_est1:AA247820	+	7.00	110.63	6.5e+03	356	AA247820	UI-HF0711	seq.f Human	
gb_est1:BB145025	+	7.00	111.17	6.0e+03	333	BB145025	BB145025	gb_est1:MW504566	+	7.00	110.63	6.5e+03	356	MW504566	UI-M-BH4-bay-d-05-0		
gb_est1:BB145025	+	7.00	111.17	6.0e+03	333	BB145025	BB145025	gb_est1:AA611403	+	7.00	110.63	6.5e+03	356	AA611403	vo51d09.r1	Barstead	
gb_gss:AO522302	+	7.00	111.17	6.0e+03	333	AO522302	BS254983	gb_est2:BI049739	+	7.00	110.63	6.5e+03	356	BI049739	CM2-GM0288-d-0201-6		
gb_est1:AAW422146	+	7.00	111.14	6.1e+03	334	AAW422146	KT210118e.KV0 Medicago	gb_est1:AJ346387	+	7.00	110.60	6.5e+03	357	AJ346387	SHR-000016-0-K19 HM		
gb_est2:BF882372	-	7.00	111.14	6.1e+03	334	BF882372	CM1-EM0192-051200-621-	gb_est1:AAW47543	+	7.00	110.60	6.5e+03	357	AAW47543	UI-M-BH1-ama-f-02-0		
gb_est1:BB446594	+	7.00	111.12	6.1e+03	335	BB446594	BB446594	gb_est1:AA592844	+	7.00	110.60	6.5e+03	357	AA592844	vo29g12.r1	Barstead	
gb_est2:BI595286	+	7.00	111.12	6.1e+03	335	BI595286	HC-L3-04B05-SKPL	gb_gss:BM261117	+	7.00	110.60	6.5e+03	357	BM261117	CH230-170P10.TV CHO		
gb_est2:ID72180	+	7.00	111.09	6.1e+03	336	ID72180	CELMK091A2R	gb_est1:AV101954	+	7.00	110.58	6.5e+03	358	AV101954	AV101954	Mus	muscul
gb_est2:FB01157	+	7.00	111.09	6.1e+03	336	FB01157	HSBA6E112	gb_est1:BM413026	+	7.00	110.58	6.5e+03	358	BM413026	EST587353	tomato	br
gb_est2:BF0159922	+	7.00	111.07	6.1e+03	337	BF0159922	BB519922	gb_est1:AM797487	+	7.00	110.56	6.5e+03	359	AM797487	CM1-UM00039-030400-1		
gb_est2:DJ17126	+	7.00	111.07	6.1e+03	337	DJ17126	CELMK0362R	gb_est2:BE399705	+	7.00	110.56	6.5e+03	359	BE399705	WHE0041	G05F90703	
gb_est2:BE341500	+	7.00	111.07	6.1e+03	337	BE341500	EST393237	gb_est1:BE399705	+	7.00	110.56	6.5e+03	359	BE399705	UI-M-BH4-bay-d-05-0		
gb_gss:AO884000	+	7.00	111.07	6.1e+03	337	AO884000	HS-5491.B2.C01.T7A RPC	gb_est1:AI747967	+	7.00	110.54	6.6e+03	360	AI747967	IB49901.x1	Zebrafis	
gb_gss:AB79585	+	7.00	111.07	6.1e+03	337	AB79585	CIT-HSP-204549.TR CIT-HS	gb_est1:AV1935056	+	7.00	110.54	6.6e+03	360	AV1935056	AV1935056	Yuj1 Kohar	
gb_est1:AA868134	+	7.00	111.05	6.1e+03	338	AA868134	vb83d06.r1	gb_gss:BM291795	+	7.00	110.51	6.6e+03	361	BM291795	CH230-30E7.TV CHOR		
gb_est1:AI047546	+	7.00	111.05	6.1e+03	338	AI047546	ub80b05.r1	gb_est1:AA038044	+	7.00	110.51	6.6e+03	361	AA038044	MR0911.r1	Soares m	
gb_est1:AB883487	+	7.00	111.05	6.1e+03	338	AB883487	AB883487	gb_est1:AA111326	+	7.00	110.51	6.6e+03	361	AA111326	mp10a08.r1	liffe Tec	
gb_est1:BB223182	+	7.00	111.05	6.1e+03	338	BB223182	BB223182	gb_est2:BF302046	+	7.00	110.51	6.6e+03	361	BF302046	BM200046	normalized	
gb_est1:AA412430	+	7.00	111.05	6.1e+03	338	AA412430	zu10c09.r1	gb_gss:AO131235	+	7.00	110.51	6.6e+03	361	AO131235	RPCT-23-289A15	TV R	
gb_est2:W08934	+	7.00	111.05	6.1e+03	338	W08934	ma60f11.r1	gb_gss:AO130257	+	7.00	110.51	6.6e+03	361	AO130257	HS-3007.A1	H07 MR C	
gb_gss:BM294582	-	7.00	111.02	6.2e+03	338	BM294582	CH230-17383.TV CHOR-2	gb_gss:AO243163	+	7.00	110.49	6.6e+03	361	AO243163	HS-2055.B2	D11 MR C	
gb_est1:AA920807	+	7.00	111.02	6.2e+03	339	AA920807	vxb6f07.r1	gb_est1:AA669894	+	7.00	110.49	6.6e+03	362	AA669894	v164b01.r1	Barstead	
gb_est2:ID17101	+	7.00	111.02	6.2e+03	339	ID17101	CELMK0742R	gb_est1:BE111051	+	7.00	110.49	6.6e+03	362	BE111051	UI-R-BJ1-auz-b-01-0		
gb_gss:AZ234544	+	7.00	111.00	6.2e+03	340	AZ234544	ATTS2859	gb_est2:BG577442	+	7.00	110.47	6.6e+03	363	BG577442	IM6-MP1Z-ADIS	006 L	
gb_gss:AZ607247	+	7.00	111.00	6.2e+03	340	AZ607247	RPCT-23-55L19.TV RCT	gb_est1:WM113990	+	7.00	110.47	6.6e+03	363	WM113990	MC6540	mouse	liver
gb_est1:AAU85896	+	7.00	110.98	6.2e+03	341	AAU85896	IM0421110R	gb_est1:AI616249	+	7.00	110.45	6.6e+03	364	AI616249	v161e03.x1	Knowles	
gb_est1:AAU008618	+	7.00	110.98	6.2e+03	341	AAU008618	mus	gb_est1:AA1616153	+	7.00	110.45	6.6e+03	364	AA1616153	mc93c05.r1	Stratagene	
gb_est1:BB16676	+	7.00	110.98	6.2e+03	341	BB16676	BB16676	gb_est2:BG225255	+	7.00	110.45	6.6e+03	364	BG225255	WHE093		

gb_gss:CNS00Y00	7.00	110.40	6.7e+03	366	AL095574	Arabidopsis thaliana g	gb_est2:BE373770	7.00	110.10	6.9e+03	380	BE373770	601226176P1	NCI_CGA
gb_gss:AO356091	7.00	110.40	6.7e+03	366	AO356091	CITR1-E1-2529B21.TR CI	gb_gss:AO373958	7.00	110.10	6.9e+03	380	AO373958	2M0187191R	Mouse 10
gb_est1:AM1096967	7.00	110.38	6.7e+03	367	AM1096967	MT11299 mouse liver c	gb_gss:AO239114	7.00	110.10	6.9e+03	380	AO239114	RPCT11-750L1.TK	RPC
gb_est2:BE4272911	7.00	110.38	6.7e+03	367	BE4272911	MTD001-1.B11990615 ID	gb_est1:AA690533	7.00	110.08	6.9e+03	381	AA690533	VC44b11.r1	Bartshead
gb_gss:BF677303	7.00	110.38	6.7e+03	367	BF677303	602087009P1 NIH.MGC.83	gb_est1:AI937107	7.00	110.08	6.9e+03	381	AI937107	wp73b10.x1	NCI_CGAP
gb_est1:AA900033	7.00	110.36	6.7e+03	367	BH302669	CH230-25G14.TJB CHORI-	gb_est1:AA680890	7.00	110.08	6.9e+03	381	AA680890	WS1-7.C05.g1.A002 W	
gb_est1:AA900033	7.00	110.36	6.7e+03	368	AA900033	UI-R-E0-dh-f-02-0-UI-S	gb_est1:AA464421	7.00	110.08	6.9e+03	381	AA464421	zx78907.x1	Soares o
gb_est1:AA545517	7.00	110.36	6.7e+03	368	AA545517	18339 MARC lBOV Bos t	gb_gss:AA256214	7.00	110.08	6.9e+03	381	AA256214	518529 MARC 3BOV Bo	
gb_est1:AA597016	7.00	110.36	6.7e+03	368	AA597016	C0194F09-3 NIA Mouse E	gb_gss:AA252023	7.00	110.08	6.9e+03	381	AA252023	RPCT-23-466P1.TVB R	
gb_est2:W48287	7.00	110.36	6.7e+03	368	W48287	VO18907.r1 Bartshead mc	gb_gss:AA982224	7.00	110.08	6.9e+03	381	AA982224	2M0262019R	Mouse 10
gb_est1:AI183783	7.00	110.34	6.7e+03	369	AI183783	qel8903.x1 Soares_fete	gb_est1:AI118237	7.00	110.06	7.0e+03	382	AI118237	qdt6e07.x1	Soares_t
gb_est2:BF731622	7.00	110.34	6.7e+03	369	BF731622	maB8907.y1 NCI_CGAP-S	gb_est1:AI128230	7.00	110.06	7.0e+03	382	AI128230	qy08e10.x1	NCI_CGAP
gb_est2:7006870	7.00	110.34	6.7e+03	369	T00681	2889 Lambda-PRI2 Arabidc	gb_est2:AM021811	7.00	110.06	7.0e+03	382	AM021811	K121b09.y1	Ascaris
gb_est2:7082167	7.00	110.34	6.7e+03	369	T02167	y939501.r1 Soares_fetal	gb_gss:AA058862	7.00	110.06	7.0e+03	382	AA058862	CIT-HSP-22966P7.TR C	
gb_est2:BE839465	7.00	110.34	6.7e+03	369	BE839465	FC3-FN0143-020800-013-	gb_gss:AA055849	7.00	110.06	7.0e+03	382	AA055849	RPCT-23-177R0.TU R	
gb_est1:AI1396740	7.00	110.31	6.7e+03	370	AI1396740	rd23b01.y1 Zebrafish W	gb_gss:AA0298503	7.00	110.06	7.0e+03	382	AA0298503	HS-2244.A1-A03.MF C	
gb_est1:AJ346762	7.00	110.31	6.7e+03	370	AJ346762	SHR-000018-0-B16 HMG/S	gb_est1:AA9239500	7.00	110.03	7.0e+03	383	AA9239500	265 P1RF2 Plnus ta	
gb_est2:BU165460	7.00	110.31	6.7e+03	370	BU165460	BU165460 full length c	gb_est1:AA925889	7.00	110.03	7.0e+03	383	AA925889	UI-R-A1-ec-h-08-0-U	
gb_est2:BE577574	7.00	110.31	6.7e+03	370	BE577574	L48-2429P3 rice plant 1	gb_est1:AV663660	7.00	110.03	7.0e+03	383	AV663660	AV663660 Bos taurus	
gb_est2:BE716508	7.00	110.31	6.7e+03	370	BE716508	CM1-HW0764-040700-299-	gb_est1:AM362988	7.00	110.03	7.0e+03	383	AM362988	RCO-CR0300-111199-0	
gb_gss:AO404814	7.00	110.31	6.7e+03	370	AO404814	GSSBRU01270 Sheard g	gb_gss:AO131186	7.00	110.03	7.0e+03	383	AO131186	HS-3028.B1-E08.MF C	
gb_est1:AA288043	7.00	110.29	6.8e+03	371	AA288043	va40d03.r1 Soares_mous	gb_est1:AI162438	7.00	110.01	7.0e+03	384	AI162438	vt04e09.x1	Bartshead
gb_est2:BG230233	7.00	110.29	6.8e+03	371	AV889043	AV889043 Nori Satoh ut	gb_est1:BI016321	7.00	110.01	7.0e+03	384	BI016321	PM3-EF0257-140301-0	
gb_est2:BI206493	7.00	110.29	6.8e+03	371	BI230233	uv04a07.x1 Soares_mous	gb_est1:AI352289	7.00	109.99	7.0e+03	385	AI352289	qr12e01.x1	Soares_t
gb_gss:AA942822	7.00	110.29	6.8e+03	371	AA942822	EST524533 CTOS Lycoper	gb_est1:AM0788271	7.00	109.99	7.0e+03	385	AM0788271	CIT-HSP-F.Lambda Zap	
gb_gss:AO244396	7.00	110.29	6.8e+03	371	AA942822	2M0203109P Mouse 10H	gb_gss:AO016366	7.00	109.99	7.0e+03	385	AO016366	CIT-HSP-23126P9.TF C	
gb_est1:AA746957	7.00	110.29	6.8e+03	372	AA248396	HS-2034.B2-D07.TV CIT	gb_gss:AA051661	7.00	109.99	7.0e+03	385	AA051661	2M0154EB03F	Mouse 10
gb_est1:BF488286	7.00	110.27	6.8e+03	372	AA746957	wt60551.f1 Aspergillus	gb_est1:AV002861	7.00	109.97	7.0e+03	386	AV002861	AO002861	Mouse 10
gb_est1:BF882501	7.00	110.27	6.8e+03	372	AA448826	zx10h07.r1 Soares_tote	gb_est1:AA409378	7.00	109.97	7.0e+03	386	AA409378	ESY90385	Mouse 7.5
gb_est2:BE406063	7.00	110.25	6.8e+03	372	BE882501	CM2-EF0193-191200-68-	gb_est2:BF884192-	7.00	109.97	7.0e+03	386	BF884192	CM2-EF0193-191200-6	
gb_est1:AA629405	7.00	110.25	6.8e+03	373	BE406063	WHR0408.f01.f012B whet	gb_gss:AA364555	7.00	109.97	7.0e+03	386	AA364555	CH230-195W2.TV CHOR	
gb_est1:BB808553	7.00	110.25	6.8e+03	373	AM629405	h154f02.x1 Soares_NFL	gb_gss:AV0873779	7.00	109.95	7.1e+03	387	AV0873779	F.rubripinus GSS sequ	
gb_est1:BB808553	7.00	110.25	6.8e+03	373	BB808553	BB808553 RIKEN full-16	gb_est1:AV004697	7.00	109.95	7.1e+03	387	AV004697	AV873779	NOI1 Satoh
gb_est1:BB872699	7.00	110.25	6.8e+03	373	BB872699	BB872699 RIKEN full-16	gb_est2:BG318365	7.00	109.95	7.1e+03	387	BG318365	NXPEV_012.G06.F NXPV	
gb_gss:AO076403	7.00	110.25	6.8e+03	373	BB872699	BB872699 RIKEN full-16	gb_gss:AO032707	7.00	109.95	7.1e+03	387	AO032707	UI-M-GCOP-bnyr.a-10-	
gb_gss:CNS00NLK	7.00	110.25	6.8e+03	373	AO076403	CIT-HSP-2359016.TR CIT	gb_gss:AA0631358	7.00	109.95	7.1e+03	387	AA0631358	RPCT-11-482B10.TV R	
gb_est1:AM046272	7.00	110.23	6.8e+03	374	AM046272	UI-M-BH1-81a-f-05-0-U	gb_est1:AA840335	7.00	109.95	7.1e+03	387	AA840335	RPCT-11-482B10.TV R	
gb_est1:AM046272	7.00	110.23	6.8e+03	374	AM046272	UI-M-BH1-81a-f-05-0-U	gb_est1:AV004697	7.00	109.95	7.1e+03	387	AV004697	AV004697 Mus muscul	
gb_est1:AA44637	7.00	110.23	6.8e+03	374	AA44637	zx85b01.r1 Soares_ovar	gb_est1:AM126511	7.00	109.93	7.1e+03	388	AM126511	UI-M-BH2-2-dq1-b-05	
gb_est1:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_est2:BG224239	7.00	109.93	7.1e+03	388	BG224239	VJ42c08.r1	Strategie
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388	BE959600	BOHPISTF BOH1 bras	
gb_est2:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132 RIKEN full-16	gb_gss:BE959600	7.00	109.93	7.1e+03	388</			

seq_name: gb_gss:A0274510

seq_documentation_block:

LOCUS A0274510 768 bp DNA linear GSS 03-NOV-1998

DEFINITION mgxb0022a11r CUGI Rice Blast BAC Library Magnaporthe grisea genomic

ACCESSION A0274510

VERSION mgxb0022a11r, DNA sequence.

KEYWORDS A0274510.1 GI:3827825

SOURCE GSS.

ORGANISM Magnaporthe grisea.

REFERENCE Magnaporthe grisea.

AUTHORS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.

1 (bases 1 to 768)

Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasnowski, M., Wang, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

TITLE Unpublished (1998)

JOURNAL Contact: Dean RA

COMMENT Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: GGAAACAGCTATGACCATG

Class: BAC ends

High quality sequence stop: 472.

Location/Qualifiers

source 1..768

/organism="Magnaporthe grisea"

/strain="70-15"

/db_xref="taxon:148305"

/clone="mgxb0022a11r"

/clone_1lb="CUGI Rice Blast BAC Library"

/lssue_type="Protoplasts"

/lab_host="E. coli DH10B"

/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

BASE COUNT 149 a 210 c 241 g 167 t 1 others

ORIGIN

alignment_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A0274510 ..

Align seg 1/1 to: A0274510 from: 1 to: 768

58 G1YTYrVAlSeRThSerLeuSerLeu 66

|||||

631 GGGTACGTAGCAGCAGCTTGTCTTG 657

seq_name: gb_gss:BH111124

seq_documentation_block:

LOCUS BH111124 104 bp DNA linear GSS 19-JUL-2001

DEFINITION RPCI-24-367M2.TJ RPCI-24 Mus musculus genomic clone RPCI-24-367M2,

ACCESSION BH111124

VERSION BH111124.1 GI:14945954

KEYWORDS GSS.

SOURCE Mus musculus.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 104)

Zhao, S., Nierman, W., Malek, J., Shatman, S., Akintet, B., Levins, M., Tsegaye, G., Geer, K., Krol, M., Shwartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-24

Unpublished (1998)

Other GSSs: RPCI-24-367M2.TJ

CONTACT: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 367 row: M column: 2

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

source 1..104

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-367M2"

/clone_1lb="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/note="Vector: pTRBAC1; Site_1: BamHI; Site_2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTRBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 17 a 24 c 35 g 28 t

ORIGIN

alignment_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BH111124 ..

Align seg 1/1 to: BH111124 from: 1 to: 104

18 ArgSerGlyGlyLeuMetProArg 25

|||||

47 AGGCTCGGGGTTTAATGCCAGG 70

seq_name: gb_est2:BI778862

seq_documentation_block:

LOCUS BI778862 137 bp mRNA linear EST 26-SEP-2001

DEFINITION EBR001_S0001_M17_R IGF Barley EBR001 library Hordeum vulgare cDNA

ACCESSION BI778862

VERSION BI778862.1 GI:15781754

KEYWORDS EST.

SOURCE Barley.

ORGANISM Hordeum vulgare

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Bakaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Trilicaceae; Hordeum.
1 (bases 1 to 137)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsey, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES
Source
Location/Qualifiers

1..137
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBR001.S0001.M17"
/clone_lib="IGF Barley EBR001 library"
/issue_type="Root, unstressed"
/lab_host="DH10b"
/note="Vector: pSPORT1; Site.1: Sal I; Site.2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old hydroponically grown
unstressed barley plants. Developed as part of the Barley
Transcriptome resources of BBSRC/SEERAD funded cereal IGF
(Investigating Gene Function) project."

BASE COUNT
ORIGIN

46 a 35 c 38 g 18 t

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x B1778662/rev ..

Align seg 1/1 to reverse of: B1778662 from: 1 to: 137

61 SerThSerLeuSerLeuArgSer 68
|||||
99 AGCACCCTCTCTCTCTCAGATCT 76

seq_name: gb_est1:BB550100

seq_documentation_block:
LOCUS BB550100 141 bp mRNA linear EST 01-AUG-2000
DEFINITION BB550100 RIKEN full-length enriched, 2 days pregnant adult female
Oviduct Mus musculus cDNA clone E230020G08 3', mRNA sequence.
ACCESSION BB550100
VERSION BB550100.1 GI:9636466
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 141)
Kono, H., Alizawa, K., Akahira, S., Akiyama, J., Arahawa, T., Carninci
P., Endo, T., Fukuda, S., Fukunishi, Y., Hata, A., Hayatsu, N.,
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M.,
Katsuyama, T., Maki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya

TITLE
JOURNAL
COMMENT

Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsunai, T., Akiyama, D., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
Source
Location/Qualifiers

1..141
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E230020G08"
/clone_lib="RIKEN full-length enriched, 2 days pregnant
adult female oviduct"
/sex="female"
/issue_type="oviduct"
/dev_stage="2 days pregnant adult"
/lab_host="DH10b"
/note="Site.1: SalI; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGGCCGACCTGACGTTTCTTTTCTTTTAA 3'], cDNA was
cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT
ORIGIN

25 a 74 c 10 g 32 t

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BB550100 ..

Align seg 1/1 to: BB550100 from: 1 to: 141

64 LeuSerLeuArgSerAlaHisLeu 71
|||||
4 CTCCTACTCGGTTCGCCGCCACCTC 27


```

seq_name: gb_est2:BF292752

seq_documentation_block:
LOCUS      BF292752              159 bp    mRNA          linear   EST 17-NOV-2000
DEFINITION WHE2202_D09_G18Z5 Aegilops speltoides anther cDNA library Aegilops
            speltoides cDNA clone WHE2202_D09_G18, mRNA sequence.
ACCESSION  BF292752
VERSION    BF292752.1  GI:11223816
KEYWORDS
SOURCE     Aegilops speltoides.
            Aegilops speltoides.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Aegilops.
REFERENCE  1 (bases 1 to 159)
AUTHORS   Akhunov,E., Anderson,O.D., Chao,S., Chin,A., Choi,D.W., Close,T.J.,
            Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Kianian,P., Izzo,G.R.,
            Miller,R., Otto,C., Rausch,C.J., Seaton,C.L., Simons,K., Tong,J.C.
            and Zhang,D.
TITLE      The structure and function of the expressed portion of the wheat
            genomes - Anther cDNA library from Aegilops speltoides
JOURNAL    Unpublished (2000)
COMMENT    Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 510535818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stragene SK primer.
            Location/Qualifiers
                1..159
                /organism="Aegilops speltoides"
                /cultivar="F2 from 2-12-4-8-1-1-(1) x P136909-12-811-(1)
                )"
                /db_xref="taxon:4573"
                /clone_1lb="Aegilops speltoides anther cDNA library"
                /tissue_type="Anther"
                /dev_stage="Premelotic anthers"
                /lab_host="E. coli SOLR"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid:
                Site_1: EcoRI, Site_2: XhoI; Plants were grown in a growth
                chamber at the University of California, Davis (Akhunov).
                Premelotic anthers were harvested, total RNA and poly(A)
                RNA were prepared, from each tissue and then pooled, a
                cDNA library was made, and the cDNA clones were in vivo
                excised to give pluescript phagemids in the TJ Close lab
                (Akhunov, Chin, Choi, Close, Fenton, Kianian, Otto, Simons
                , Zhang) at the University of California, Riverside.
                Plasmid DNA preparations and DNA sequencing were performed
                in the OD Anderson lab (all other authors)."
```

```

BASE COUNT      34 a          29 c          45 g          51 t
ORIGIN
```

```

alignment_scores:
    Quality:      8.00          Length:      8
    Ratio:        1.000         Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
```

```
US-09-528-682-1 x BF292752 ..
```

```
Align seg 1/1 to: BF292752 from: 1 to: 159
```

```

61 SerthrSerleuSerleuArgSer 68
|||||
34 TCGACAGAGCTGTCTCGCAAGT 57
seq_name: gb_est2:BE517249
```

```

seq_documentation_block:
LOCUS      BE517249              163 bp    mRNA          linear   EST 08-AUG-2000
DEFINITION WHE0616_G08_N16Z4 wheat ABA-treated embryo cDNA library Triticum
            aestivum cDNA clone WHE0616_G08_N16, mRNA sequence.
ACCESSION  BE517249
VERSION    BE517249.1  GI:9741279
KEYWORDS
SOURCE     bread wheat.
            Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 163)
AUTHORS   Anderson,O.D., Chao,S., Han,P.S., Hsia,C.C., Johnson,R.R., Kang,Y.,
            Izzo,G.R., Miller,R., Rausch,C.J., Seaton,C.L., Tong,J.C., Verhey
            ,S.D. and Walker-Simmons,M.K.
TITLE      The structure and function of the expressed portion of the wheat
            genomes - ABA-treated embryo library
JOURNAL    Unpublished (2000)
COMMENT    Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Clontech Matchmaker 3' AD primer.
            Location/Qualifiers
                1..163
                /organism="Triticum aestivum"
                /cultivar="Brevor (soft, white, winter, common wheat)"
                /db_xref="taxon:4565"
                /clone_1lb="WHE0616_G08_N16"
                /clone_1lb="Wheat ABA-treated embryo cDNA library"
                /tissue_type="Seed embryo"
                /dev_stage="Seed dormant seeds"
                /lab_host="E. coli DH12S"
                /note="Vector: pGAD10; Site_1: EcoRI; Site_2: XhoI;
                Embryos were cut from mature, dormant seeds and limbed in
                25 microm ABA (abscisic acid) in 5 mm MMS buffer, pH 5.7,
                for 12 hr at 22 C. The tissue, total RNA, and poly(A)
                RNA were prepared by Steven Verhey in M.K.
                Walker-Simmons' lab (USDA-ARS, Washington State Univ.,
                Pullman, Washington 99164-6420). A cDNA library was made
                by Clontech using a combination of random and oligo dT
                primers. Library was plated and archived by Russell
                Johnson (Colby College, ME/Walker-Simmons' lab). Plasmid
                DNA preparations and DNA sequencing were performed in the
                OD Anderson lab (all other authors)."
```

```

BASE COUNT      39 a          36 c          54 g          34 t
ORIGIN
```

```

alignment_scores:
    Quality:      8.00          Length:      8
    Ratio:        1.000         Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
```

```
US-09-528-682-1 x BE517249 ..
```

```
Align seg 1/1 to: BE517249 from: 1 to: 163
```

```

61 SerthrSerleuSerleuArgSer 68
|||||
93 TCGACAGAGCTGTCTCGCAAGT 116
seq_name: gb_est1:AA637879
seq_documentation_block:
```


LOCUS AA637879 169 bp mRNA linear EST 22-OCT-1997
 DEFINITION vir29h03.r1 Barstead mouse myotubes MRLB5 Mus musculus cDNA clone
 IMAGE:1122101 5' similar to TR:G1020151 G1020151 CLP36. ;, mRNA
 sequence.
 ACCESSION AA637879.1 GI:2561467
 VERSION AA637879.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 169)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Gettel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:611437
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 EF from Amersham
 High quality sequence stop: 1.
 FEATURES
 source
 1..169
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C3H"
 /db_xref="taxon:10090"
 /clone="IMAGE:1122101"
 /clone_lib="Barstead mouse myotubes MRLB5"
 /cell_line="C2C12"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCGATCGATCGAGCGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [AATTCGATCCTTCG], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified p7T73 vector.
 Library constructed by Bob Barstead. The C2C12 cell line
 (available from ATCC, catalog # CRL-1772) differentiates
 rapidly, forming contractile myotubes and producing
 characteristic muscle proteins."
 BASE COUNT 30 a 49 c 45 g 45 t
 ORIGIN
 alignment_scores: Length: 8
 Quality: 1.000 Gaps: 0
 Ratio: 1.000
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x AA637879 ..
 Align seg 1/1 to: AA637879 from: 1 to: 169
 62 ThrSerLeuSerLeuArgSerAla 69
 ||||||||||||||||||||
 51 ACTTCACCTCAGCGCTCGGTCAGCA 74
 seq_name: gb_est2:BM373106

seq_documentation_block:
 LOCUS BM373106 232 bp mRNA linear EST 10-JAN-2002
 DEFINITION EBMa04_S0003_C02_R IGF Barley EBMa04 library Hordeum vulgare cDNA
 clone EBMa04_S0003_C02 5', mRNA sequence.
 ACCESSION BM373106
 VERSION BM373106.1
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 232)
 AUTHORS Redway, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
 Ramsay, L., Hare, G., Marshall, D.F.M. and Maugh, R.
 TITLE Development of Barley Transcriptome Resources
 JOURNAL Unpublished (2001)
 COMMENT Contact: Maugh R
 Unit of Genomics
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: rwaugh@scri.sari.ac.uk
 All sequence has a Phred quality score of 20 or over
 Seq primer: M13 reverse.
 FEATURES
 source
 1..232
 Location/Qualifiers
 /organism="Hordeum vulgare"
 /cultivar="Optic"
 /db_xref="taxon:4513"
 /clone="EBMa04_S0003_C02"
 /clone_lib="IGF Barley EBMa04 library"
 /tissue_type="Maternal tissue"
 /dev_stage="10 days post anthesis"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from maternal tissue dissected from developing
 grains (10 days post anthesis) in glasshouse grown barley
 plants. Developed as part of the barley transcriptome
 resources of BBSRC/NERAC funded cereal IGF (Investigating
 Gene Function) project."
 BASE COUNT 32 a 82 c 77 g 41 t
 ORIGIN
 alignment_scores: Length: 8
 Quality: 1.000 Gaps: 0
 Ratio: 1.000
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x BM373106/rev ..
 Align seg 1/1 to reverse of: BM373106 from: 1 to: 232
 61 SerThrSerLeuSerLeuArgSer 68
 ||||||||||||||||||||
 45 AGCACCTCTCTCTCTCAGATCT 22
 seq_name: gb_est1:AW202162
 seq_documentation_block:
 LOCUS AW202162 254 bp mRNA linear EST 02-DEC-2001
 DEFINITION s112f04.y1 Gm-c1027 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1027-2072 5' similar to TR:Q92V18 Q92V18 T6A23.14 PROTEIN. ;,
 mRNA sequence.
 ACCESSION AW202162
 VERSION AW202162.1
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max

circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-CO library. This procedure has been previously described (Bonald, Lennon and Soares, Genome Research 6: 791-806, 1996).

BASE COUNT 66 a 50 c 66 g 76 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA998803/rev ..

Align seg 1/1 to reverse of: AA998803 from: 1 to: 258

60 ValSerThrSerLeuSerLeuArg 67
|||||
152 GTATCTACTCTCTCTCTCTAAGG 129

seq_name: gb_gss:A2788524

seq_documentation_block:

LOCUS A2788524 263 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0035K08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0035K08 R, DNA sequence.

ACCESSION A2788524
VERSION A2788524.1 GI:12928413

KEYWORDS GSS.

SOURCE

ORGANISM

Mus musculus.
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 263)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

Unpublished (2000)

JOURNAL

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0035 row: K column: 08

Seq primer: CACACAGGAACACGATGACG

Class: plasmid ends

High quality sequence strip: 263.

Location/Qualifiers

FEATURES

source

1.263

/organism="Mus musculus"

/strain="C57Bl/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0035K08"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD42ny; Purified genomic DNA from M.
musculus C57Bl/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 47 a 81 c 78 g 57 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A2788524 ..

Align seg 1/1 to: A2788524 from: 1 to: 263

69 AAlaHStleuAlaGlyGlnSerIle 76
|||||
161 GCTCCTCCTGCGGAGCAGACATT 184

seq_name: gb_gss:AQ911978

seq_documentation_block:

LOCUS AQ911978 269 bp DNA linear GSS 25-MAY-2001
DEFINITION LMAJFV1_In09B09.y1 Leishmania major FV1 random genomic library
Leishmania major genomic clone LMAJFV1_In09B09 5', DNA sequence.

ACCESSION AQ911978
VERSION AQ911978.1 GI:6508494

KEYWORDS GSS.

SOURCE

ORGANISM

Leishmania major.
Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.

1 (bases 1 to 269)

JOURNAL

Akopyants, N.S., Clifton, S.W., Martin, J., Pape, D., Wylie, T., Li, L.,
Kissinger, J., Roos, D.S., Marra, M., Hillier, L., Chinwalla, A.,
Bistain, A., Schmitt, A., Person, B., Theising, B., Ritter, E., Ronko,
I., Bennett, J., Cole, R., Underwood, R., Cardenas, M., Gibbons, M.,
Harvey, N., McCann, R., Tsagaris, R., Williams, T., Jackson, Y.,
Bowers, Y., Swaller, T., Waterston, R., Wilson, R. and Beverley, S.M.
A survey of the Leishmania major Friedlin strain Y1 genome by
shotgun sequencing: a resource for DNA microarrays and expression
profiling

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

21192569

Other GSSs: In09B09.x1

Contact: Akopyants, NS / Beverley, SM

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Library construction: Natalia S. Akopyants, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

If using this information please cite:

N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major

Friedlin strain Y1 genome by shotgun sequencing' and the Washington

University Genome Sequencing Center for information on obtaining

clone material please contact: Natalia S. Akopyants Ph.D.

(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.


```

/dev stage="20 week-post conception fetus"
/lab_host="TDH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLs library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5],
AACGGAAGATTAATTAACATCTTTTCTTTTCTTTTCTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."

```

```
alignment_scores:
  Quality: 8.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 8
  Gaps: 0
  Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-1 x W90748/rev ..
```

Align seg 1/1 to reverse of: W90748 from: 1 to: 309

220 ArgGlnIlePheSerAspTyrGln 227
 |||
 69 AGACAATCTTTCTGATTATCAA 46

seq_name: gb_est2:BI050522

LOCUS	312 bp	mRNA	linear	EST 15-JUN-2001
BI050522				
DEFINITION	PM2-GN0015-060101-008-f04	GN0015	Homo sapiens CDNA,	mRNA sequence.
DESCRIPTION				

VERSION BI050522.1 GI:14458052
KEYWORDS EST.

ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Vertjovski-Almeida, S., Britones, M. R.,
1 (bases 1 to 312)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Nagai, M. A. da Silva, W. Jr., Zazo, M. A., Bordini, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
ABSTRACT 00000000

COMMENT
Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PM2&t2=PM2-GN0015-060101-008-104&t3=2001-01-10&t4=1>)

seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 226

FEATURES	Location/Qualifiers
source	1. .312

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

```

/clone_lib="GN0015"
/dev_stage="Adult"
/note="Organ: Placenta.normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
58 a 98 c 87 g 69 t

```

```
alignment_scores:
  quality: 8.00      Length: 8
  ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block: 00000000-0000-0000-0000-000000000000

Align seg 1/1 to reverse of: BI050522 from: 1 to: 312

111 GlnGluValSerAlaLeuGlyGly 118
 |||
 176 CAGGAGGTCACGCACTGGGGCC 153

seq_name: gb_est2:BI781014

LOCUS	B1781014	312 bp	mRNA	linear	EST 26-SEP-2001
DEFINITION	EBma03_S0001_c14_R	IGF Barley	EBma03 library	Hordeum vulgare	CDNA
DESCRIPTION	clone EBma03_S0001_c14 5', mRNA sequence.				

VERSION BI781014.1 GI:15783906
KEYWORDS EST.

Source	Organism	Library
	<i>Hordeum</i>	
	Eukaryote	

REFERENCE
1 (bases 1 to 312)
; Triticaceae; Hordeum.

AUTHORS	Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudge, Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
TITLE	Development of Barley Transcriptome Resources

JOURNAL Unpublished (2001),
COMMENT Contact: Waugh R

Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK

Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk

Seq primer: M13 reverse.

```

FEATURES
source
1. .312
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="Optic"

```

```
/db_xref="taxon:4513"  
/clone="EBma03_SQ001_
```

```

/clone_lib="IGF Barley Emma03 library"
/tissue_type="Maternal tissue"
/dev_stage="8 days post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;

```

Non-normalised library, directionally cloned into pSPORT1. Derived from maternal tissue dissected from developing grains (8 days post anthesis) in glasshouse grown barley

plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating

BASE COUNT	83 a	54 c	88 g	87 t
------------	------	------	------	------

ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BT781014 ..

Align seg 1/1 to: BT781014 from: 1 to: 312

61 SerThrsSerLeuSerLeuArgSer 68
|||||
77 TCGACAGCTGTCTCTCGCAGAGT 100

seq_name: gb_est2:BM101253

seq_documentation_block:

LOCUS BM101253 320 bp mRNA linear EST 21-NOV-2001
DEFINITION EBP101_SQ003_F08.R IGF Barley EBP101 library Hordeum vulgare cDNA
clone EBP101_SQ003_F08 5', mRNA sequence.

ACCESSION BM101253 GI:17032321

VERSION EST.

KEYWORDS

SOURCE

ORGANISM barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 320)

REFERENCE

AUTHORS

Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)

JOURNAL

COMMENT

Contact: Waugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse

FEATURES

source

Location/Qualifiers
1..320
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBP101_SQ003_F08"
/clone_lib="IGF Barley EBP101 library"
/tissue_type="Pistils"
/dev_stage="24 hours post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from pistils dissected from developing grains (24
hours post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."

BASE COUNT 68 a 57 c 98 g 97 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BM101253 ..

Align seg 1/1 to: BM101253 from: 1 to: 320

61 SerThrsSerLeuSerLeuArgSer 68
|||||

84 TCGACAGCTGTCTCTCGCAGAGT 107

seq_name: gb_est2:BF778513

seq_documentation_block:

LOCUS BF778513 323 bp mRNA linear EST 12-JAN-2001
DEFINITION NXSI_085_F03_F NXSI (Nsif Xylem Side wood Inclined) Pinus taeda cDNA
clone NXSI_085_F03 5', mRNA sequence.

ACCESSION BF778513 GI:12126413

KEYWORDS

SOURCE

ORGANISM

loblolly pine.
Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 323)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Johnson, Archur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: a.johnson@unity.ncsu.edu
Seq primer: T3.

FEATURES

source

Location/Qualifiers
1..323
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXSI_085_F03"
/clone_lib="NXSI (Nsif Xylem Side wood Inclined)"
/tissue_type="Xylem"
/cell_type="Side"
/dev_stage="Juvenile"
/lab_host="XL1-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form side
wood by bending to a 45 degree angle and tying them to the
ground. Differentiating xylem was harvested from the sides
of the inclined stems, and a mixture of all three
genotypes was used for the library. oligo-dT primed cDNA
was directionally cloned into the EcoRI-XhoI Bluescript SK
vector arms. NOTE: The sequences contain a 'cDNA adapter'
between the EcoRI site and the start of the EST. The
adapter sequence is 'ATTCCGACGAG'."

BASE COUNT 59 a 84 c 82 g 89 t 9 others
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BF778513 ..

Align seg 1/1 to: BF778513 from: 1 to: 323

14 AspgluileysArgSerGly 21
|||||

216 GATGAGATCAAGCGGAGCGGG 239

seq_name: gb_est2:BM368647


```

seq_documentation_block:
LOCUS      BM368647          333 bp      mRNA      linear      EST 10-JAN-2002
DEFINITION EBem08_SQ004_D10_R IGF Barley EBem08 library Hordeum vulgare cDNA
ACCESSION  BM368647
VERSION     BM368647.1  GI:18112037
KEYWORDS    EST.
SOURCE      Barley.
ORGANISM    Hordeum vulgare
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
             ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 333)
AUTHORS     Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardie,L.,
             Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE       Development of Barley Transcriptome Resources
JOURNAL     Unpublished (2001)
COMMENT     Contact: Waugh R
             Unit of Genomics
             Scottish Crop Research Institute
             Invergowrie, Dundee, DD2 5DA, Scotland, UK
             Tel: 00 44 1382 562731
             Fax: 00 44 1382 562426
             Email: rwaugh@scri.sari.ac.uk
             All sequence has a Phred quality score of 20 or over
             Seq primer: M13 reverse.
FEATURES
             Location/Qualifiers
             source          1..333
                             /organism="Hordeum vulgare"
                             /cultivar="Optic"
                             /db_xref="taxon:4513"
                             /clone="EBem08_SQ004_D10"
                             /clone_lib="IGF Barley EBem08 library"
                             /tissue_type="Embryo"
                             /dev_stage="40 days post anthesis"
                             /lab_host="DH10B"
                             /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                             Non-normalised library, directionally cloned into pSPORT1.
                             Derived from embryos dissected from developing grains (40
                             days post anthesis) in glasshouse grown barley plants.
                             Developed as part of the barley transcriptome resources of
                             BBSRC/SEERD funded cereal IGF (Investigating Gene
                             Function) project."
BASE COUNT      80 a          58 c          99 g          96 t
ORIGIN
alignment_scores:
Quality:      8.00      Length:      8
Ratio:        1.000     Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x BM368647 ..
Align seg 1/1 to: BM368647 from: 1 to: 333
61 SerThrsSerLeuSerLeuArgSer 68
|||||
109 TCACACAAAGCTTGTCTCTCGCAAGT 132
seq_name: gb_gss:AZ094117
seq_documentation_block:
LOCUS      AZ094117          346 bp      DNA      linear      GSS 08-MAY-2000
DEFINITION RPCI-23-448B16.T1 RPCI-23 Mus musculus genomic clone RPCI-23-448B16
ACCESSION  AZ094117
VERSION     AZ094117.1  GI:77361610
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE     1 (bases 1 to 346)
AUTHORS       Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akintet
              ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
              and Fraser,C.M.
TITLE         Mouse BAC End Sequences from Library RPCI-23
JOURNAL       Unpublished (1999)
COMMENT       Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@ligr.org
              Clones are derived from the mouse BAC library RPCI-23. For BAC
              library availability, please contact Pieter de Jong
              (pieter@dejong.med.buffalo.edu). Clones may be purchased from
              BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
              or from Resea ch Genetics (info@resgen.com). BAC end page:
              http://www.ligr.org/tldb/bac_ends/mouse/bac_end_intro.html
              Plate: 448 row: B column: 16
              Seq primer: SP6
              Class: BAC ends.
FEATURES
             Location/Qualifiers
             source          1..346
                             /organism="Mus musculus"
                             /strain="C57BL/6J"
                             /db_xref="taxon:10090"
                             /clone="RPCI-23-448B16"
                             /clone_lib="RPCI-23"
                             /sex="Female"
                             /lab_host="DH10B"
                             /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
                             EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                             brain genomic DNA was isolated and partially digested
                             with a combination of EcoRI and EcoRI Methylase. Size
                             selected DNA was cloned into the pBAC3.6 vector at the
                             EcoRI sites. The ligation products were transformed into
                             DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      94 a          70 c          97 g          85 t
ORIGIN
alignment_scores:
Quality:      8.00      Length:      8
Ratio:        1.000     Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AZ094117/rev ..
Align seg 1/1 to reverse of: AZ094117 from: 1 to: 346
61 SerThrsSerLeuSerLeuArgSer 68
|||||
104 AGCACCCTCTCTAGCGCTCC 81
seq_name: gb_gss:AQ123154
seq_documentation_block:
LOCUS      AQ123154          355 bp      DNA      linear      GSS 22-SEP-1998
DEFINITION HS_3095_B1_F12_MR CIT Approved Human Genomic Sperm Library D Homo
              sapiens genomic clone Plate=3095 Col=23 Row=L, DNA sequence.
ACCESSION  AQ123154
VERSION     AQ123154.1  GI:3500320
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 355)
AUTHORS       Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
              Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

```


TITLE
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL
MEDLINE
99380588
COMMENT
Contact: Mahatras GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3095 row: L column: 23
Class: BAC ends
High quality sequence stop: 355.
Location/Qualifiers

FEATURES
source
1..355
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate:3095 Col=23 Row=L"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT 145 a 83 c 57 g 70 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ123154/rev ..
Align seg 1/1 to reverse of: AQ123154 from: 1 to: 355

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
72 GTTTCGCTTTCGAGGATCCCA 49

seq_name: gb_estc2:BI777758

seq_documentation_block:

LOCUS BI777758 356 bp mRNA linear EST 26-SEP-2001
DEFINITION EBR008_SQ001_003_R IGF Barley EBR008 library Hordeum vulgare cDNA
clone EBR008_SQ001_003 5', mRNA sequence.
ACCESSION BI777758
VERSION BI777758.1 GI:15780650
KEYWORDS EST.

SOURCE
ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 356)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsey, L., Machray, G., Marshall, D. F. M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
source

Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers
1..356

/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBR008_SQ001_003"
/clone_lib="IGF Barley EBR008 library"
/issue_type="Drought stressed root"
/lab_host="DH10B"
/note="Vector: pSPORT1; site: 1: Sal I; site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old drought stressed barley
plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal IGF (investigating
Gene Function) project."
BASE COUNT 74 a 115 c 105 g 62 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BI777758/rev ..
Align seg 1/1 to reverse of: BI777758 from: 1 to: 356

61 SerThrSerLeuSerLeuArgSer 68
|||||
95 AGCACCTCTCTCTCTCAGATCT 72

seq_name: gb_estc2:BM441268

seq_documentation_block:

LOCUS BM441268 357 bp mRNA linear EST 01-FEB-2002
DEFINITION EBed02_SQ002_006_R IGF Barley EBed02 library Hordeum vulgare cDNA
clone EBed02_SQ002_006 5', mRNA sequence.
ACCESSION BM441268
VERSION BM441268.1 GI:18472043
KEYWORDS EST.

SOURCE
ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 357)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsey, L., Machray, G., Marshall, D. F. M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
source

Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers
1..357
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBed02_SQ002_006"
/clone_lib="IGF Barley EBed02 library"
/issue_type="Endosperm"
/dev_stage="8 days post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; site: 1: Sal I; site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from endosperm tissue dissected from developing
grains (8 days post anthesis) in glasshouse grown barley

plants. Developed as part of the barley transcriptome
resources of BSRG/SEAD funded cereal IGF (Investigating
Gene Function) project."

BASE COUNT 61 a 123 c 112 g 61 t

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BM441268/rev ..

Align seg 1/1 to reverse of: BM441268 from: 1 to: 357

61 SerTherSerleuSerleuArgSer 68
|||||
131 AGCACCCTCTCTCTCTCAGATCT 108

seq_name: gb_est2:c60468

seq_documentation_block:

LOCUS C60468 360 bp mRNA linear EST 22-SEP-1997
DEFINITION C60468 yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk208a11 5', mRNA sequence.

ACCESSION C60468
VERSION C60468.1 GI:2419173

KEYWORDS EST.
SOURCE Caenorhabditis elegans.

ORGANISM

Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae

; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE

1 (bases 1 to 360)
Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano

AUTHORS

TITLE

JOURNAL

COMMENT

Expression map of the C.elegans genome
Unpublished (1996)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

/clone="yk208a11"
/clone_lib="yuji Kohara unpublished CDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 106 a 72 c 78 g 104 t

ORIGIN

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

106 a 72 c 78 g 104 t

seq_name: gb_est2:c64952

seq_documentation_block:

LOCUS C64952 360 bp mRNA linear EST 22-SEP-1997
DEFINITION C64952 yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk397d10 5', mRNA sequence.

ACCESSION C64952
VERSION C64952.1 GI:2423657

KEYWORDS EST.
SOURCE Caenorhabditis elegans.

ORGANISM

Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae

; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE

1 (bases 1 to 360)
Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano

AUTHORS

TITLE

JOURNAL

COMMENT

Expression map of the C.elegans genome
Unpublished (1996)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

/clone="yk397d10"
/clone_lib="yuji Kohara unpublished CDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

105 a 71 c 77 g 107 t

Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

source

1..360

Location/Qualifiers

/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

/clone="Yk59g12"

/clone_lib="Yuji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 99 a 71 c 81 g 97 t 12 others
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x D69091 ..

Align seg 1/1 to: D69091 from: 1 to: 360

137 GUARGLEUHSARGASARGIIL 144
|||||
230 GAAAGGCTGCACAGAACCTGAG 253

seq_name: gb_est1:AW740520

seq_documentation_block:

LOCUS AW740520 363 bp mRNA linear EST 27-APR-2000

DEFINITION BR110477 Blomphalaria glabrata (BS-90)-unexposed lambda zap library

Blomphalaria glabrata cDNA clone RBG1G25TR, mRNA sequence.

ACCESSION AW740520

VERSION AW740520.1 GI:7651613

KEYWORDS EST.

SOURCE bloodfluke planorb.

ORGANISM Blomphalaria glabrata

Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora

; Planorbidae; Blomphalaria.

1 (bases 1 to 363)

Raghavan, N., Miller, A., Gardner, M., Kerlavage, A.R., Fitzgerald, P.C.

and after exposure to miracidia

Unpublished (2000)

JOURNAL Contact: Raghavan N

COMMENT Biomedical Research Institute

1211 Parklawn Dr., Rockville, MD 20852, USA

Tel: 301-881-3300 ext.128

Fax: 301-770-4756

Email: nkrhelix.nih.gov, snailsrule@aol.com.

FEATURES

source

1..363

Location/Qualifiers

/organism="Blomphalaria glabrata"

/strain="BS-90"

/db_xref="taxon:6526"

/clone="RBG1G25TR"

/clone_lib="Blomphalaria glabrata (BS-90)-unexposed lambda

zap library"

/sex="hermaphrodite"

/cell_type="Hemocyte"

/lab_host="Laboratory host"

/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:

XhoI; Total RNA was isolated from the hemocytes of

unexposed Blomphalaria glabrata (BS-90) snails and first

strand cDNA synthesized using an oligo-dT primer-linker
(XhoI). Second strand synthesis was followed by the
ligation of EcoRI adaptors. Following digestion with XhoI,
the completed, directional cDNA was cloned into Uni-ZAP
XR phagemid vector by Stratagene.

BASE COUNT 109 a 77 c 71 g 106 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AW740520/rev ..

Align seg 1/1 to reverse of: AW740520 from: 1 to: 363

61 SerThSerLeuSerLeuArgSer 68

|||||

seq_name: gb_gss:AO099964

seq_documentation_block:

LOCUS AO099964 369 bp DNA linear GSS 27-NOV-1998

DEFINITION HS_3054_B1_B07_MF CIT Approved Human Genomic Sperm Library D Homo

sapiens genomic clone Plate=3054 Col=13 Row=D, DNA sequence.

ACCESSION AO099964

VERSION AO099964.1 GI:3470993

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 369)

Mahalras, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,

Keller, A., Shaker, R., Furlong, V., Young, V., Zhao, S., Adams, M.D. and

Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

JOURNAL Contact: Mahalras GG, Wallace JC, Hood L

MEDLINE High Throughput Sequencing Center

COMMENT University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3054 Row: D Column: 13

Class: BAC ends

High quality sequence stop: 369.

FEATURES

source

1..369

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate=3054 Col=13 Row=D"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones 1n

E-Coli DH10B

BASE COUNT 78 a 100 c 64 g 127 t

ORIGIN

alignment_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A0099964 ..
Align seg 1/1 to: A0099964 from: 1 to: 369

64 LeuSerLeuArgSerAlaHisLeu 71
|||||
167 CTTTCCTGAGGCTGCTCACCTC 190

seq_name: gb_gss:A0544348

seq_documentation_block:
LOCUS A0544348 370 bp DNA linear GSS 28-MAY-1999
DEFINITION CITBI-E1-2651G15.TR CITBI-E1 Homo sapiens genomic clone 2651G15,
DNA sequence.
ACCESSION A0544348
VERSION A0544348.1 GI:4903423
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other-GSSs: CITBI-E1-2651G15.TF
Contact: Shaying Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2651G15"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelOBAC11, Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 79 a 65 c 80 g 146 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A0544348 ..
Align seg 1/1 to: A0544348 from: 1 to: 370

113 ValSerAlaLeuGlyIlePro 120
|||||
305 GTATCTGCTTTAGGAGCATCCCA 328

seq_name: gb_est2:H78351

seq_documentation_block:
LOCUS H78351 372 bp mRNA linear EST 09-NOV-1995

DEFINITION yu79f12.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:240047 5', mRNA sequence.
ACCESSION H78351
VERSION H78351.1 GI:1056440
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 372)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kuwaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Ritkin,L., Rohlfing,T., Soares,M., Tan,F., Treaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 692
High quality sequence stops: 236
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 692 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 236.

FEATURES
source
1..372
Location/Qualifiers
1..372
/organism="Homo sapiens"
/db_xref="GDB:3789000"
/db_xref="taxon:9606"
/clone="IMAGE:240047"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15', ACTCGAGAGATTATATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 76 a 96 c 93 g 103 t 4 others
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x H78351 ..
Align seg 1/1 to: H78351 from: 1 to: 372

113 ValSerAlaLeuGlyIlePro 120
|||||
7 GTGTCTGCATTAGGGGAGCATCCCA 30

seq_name: gb_gss:A2081963

seq_documentation_block:
LOCUS A2081963 372 bp DNA linear GSS 08-JAN-2001
DEFINITION UP-566-2E.T7 RPC111 Human Male BAC Library Homo sapiens genomic
clone 566-2E, DNA sequence.


```

ACCESSION      AZ081963
VERSION        AZ081963.1  GI:7709217
KEYWORDS       GSS.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS        Cheung,V.G., Daitymple,H.L., Narasimhan,S., Watts,J., Schuler,G.,
                Raap,A.K., Morley,M. and Bruzel,A.
TITLE          A resource of mapped human bacterial artificial chromosome clones
JOURNAL        Genome Res. 9 (10), 989-993 (1999)
MEDLINE        99455100
COMMENT        Contact: Arcaro MA, Morley M, Burdick J, Cheung VC
                Department of Pediatrics
                University of Pennsylvania
                3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA
                Tel: 215 590 2664
                Fax: 215 590 3709
                Email: mlennox@mail.med.upenn.edu
                Plate: 566 row: E column: 2
                Seq primer: 77
                Class: BAC ends.

FEATURES
  source        1..372
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_1b="RPC111 Human Male BAC Library"
                /sex="Male"
                /cell_type="Lymphocytes"
                /note="Vector: pBAC3.6; RPC111 Human Male BAC Library"

BASE COUNT     90 a 69 c 46 g 167 t
ORIGIN

alignment_scores:
  Quality:      8.00      Length:      8
  Ratio:        1.000     Gaps:        0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AZ081963/rev ..

Align seg 1/1 to reverse of: AZ081963 from: 1 to: 372

217 LysValLysArgGlnIlePheSer 224
|||||
311 AAGCTAAAGACAAATATTTTCT 288

seq_name: gb_gss:BH462450

seq_documentation_block:
LOCUS          BH462450 381 bp DNA linear GSS 13-DEC-2001
DEFINITION    BOKH235TF BOKH Brassica oleracea genomic clone BOKH235, DNA
sequence.
ACCESSION     BH462450
VERSION       BH462450.1 GI:17654289
KEYWORDS      GSS.
SOURCE        Brassica oleracea.
ORGANISM      Brassica oleracea
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 381)
AUTHORS        Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other_GSSs: BOKH235TR
                Contact: Chris Town
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA.
                Tel: 301-938-3523

Fax: 301-938-0208
Email: ctown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 77
Class: sheared ends.

FEATURES
  source        1..381
                /organism="Brassica oleracea"
                /strain="T01000DH3"
                /db_xref="taxon:3712"
                /clone="BOKH235"
                /clone_1b="BOKH"
                /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
                genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT     135 a 73 c 82 g 91 t
ORIGIN

alignment_scores:
  Quality:      8.00      Length:      8
  Ratio:        1.000     Gaps:        0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BH462450/rev ..

Align seg 1/1 to reverse of: BH462450 from: 1 to: 381

60 ValSerThrSerLeuSerLeuArg 67
|||||
82 GTATCACTTCTCTCAGCCGTCAG 59

seq_name: gb_est2:BE521751

seq_documentation_block:
LOCUS          BE521751 382 bp mRNA linear EST 19-MAR-2001
DEFINITION    M21C8STM Arabidopsis developing seed Arabidopsis thaliana cDNA
clone M21C8 5', mRNA sequence.
ACCESSION     BE521751
VERSION       BE521751.1 GI:9779729
KEYWORDS      EST.
SOURCE        thale cress.
ORGANISM      Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE      1 (bases 1 to 382)
AUTHORS        White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
                Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
TITLE          A new set of Arabidopsis expressed sequence tags from developing
                seeds. The metabolic pathway from carbohydrates to seed oil
                plant Physiol. 124 (4), 1582-1594 (2000)
JOURNAL        20567808
MEDLINE
COMMENT        Contact: Benning, C
                Dept. of Biochemistry & Molecular Biology
                Michigan State University
                224 Biochemistry, Michigan State University, East Lansing, MI 48824
                , USA
                Tel: 517 355 1609
                Fax: 517 353 9334
                Email: benning@msu.edu
                Michigan State University DNA Sequencing Facility Arabidopsis
                Biological Resource Center, The Ohio State University, 309 Botany &
                Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
                6142920603 TEL: 6142929371.

FEATURES
  source        1..382
                /organism="Arabidopsis thaliana"
                /strain="Columbia"
                /db_xref="taxon:3702"
                /clone="M21C8"
                /clone_1b="Arabidopsis developing seed"
                /tissue_type="seed"

```



```

/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/notes="Organ: Developing seed; Vector: pBluescript SK-
Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      119 a      90 c      87 g      86 t
ORIGIN

```

```
alignment_scores:
  Quality: 8.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 8
  Gaps: 0
  Percent Identity: 100.0000
```

```
alignment_block:
US-09-528-682-1 x BE521751 .
```

Align seg 1/1 to: BE521751 from: 1 to: 382

204 GlnAsnLeuSerThrIleTyrLeu 211
 |||
 336 CAGAACTTGTCACACATCTATTTA 359

seq_name: gb_est2:BF010577

LOCUS	BF010577	390 bp	mRNA	linear	EST 06-OCT-2000
DEFINITION	NXCI_086_c11_F	NXCI (nsf xylem	compression	wood	inclined) plus
DESCRIPTION	taeda cDNA clone NXCI_086_c11 5', mRNA sequence.				

ACCESSION	BF010577
VERSION	BF010577.1
	GI:10710851

ORGANISM

REFERENCE
1 (bases 1 to 390)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Coniferopsida, Coniferales; Pinaceae; Pinus; Pinus

TITLE	Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL	unpublished (2000)
COMMENT	Contact: Johnson, Arthur

North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801

Seq primer: T3

FEATURES	
source	Location/Qualifiers
	1. .390

```

/organism="Pinus taeda"
/strain="Coastal plain loblobly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXCI_086_C11"
/clone_lib="NXCI (Nsf Xylem Compression wood Inclined)"
/tissue_type="Xylem"
/cell_type="Compression"
/dev_stage="Juvenile"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form
compression wood by bending to a 45 degree angle and tying
them to the ground. Differentiating xylem was harvested
from the bottoms of the inclined stems, and a mixture of
all three genotypes was used for the library. oligo-dT
primed cDNA was directionally cloned into the EcoRI-XhoI
Bluescript SK vector arms. NOTE: The sequences contain a
'cDNA adapter' between the EcoRI site and the start of the
EST. The adapter sequence is 'AATTCGACACAG'."

```

```

BASE COUNT
ORIGIN
alignment_scores:

```

Quality:	8.00	Length:	8
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-528-682-1 x BF010577 .

Align seg 1/1 to: BF010577 from: 1 to: 390

14 AspGIuIleLysArgSerGlyGly 21
 |||||
 218 GATGAGATCAAGCGGAGCGGCGG 241

seq_name: gb_est2:W78167

```
seq_documentation_block:
  docnum: 170167
```

LOCUS	393 bp	mRNA	linear	EST 1 / OCT-1998
DEFINITION	z79g08.s1 Soares fetal heart.NbHH19W Homo sapiens cdna clone			
IMAGE:	346910 3', mRNA sequence.			

VERSION W78167.1 GI:1388701

SOURCE	human
...	...

ORGANIS

ORGANISM	REFERENCE
<i>Homo sapiens</i>	Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
1 (bases 1 to 393)	

JOURNAL

COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 671 Std Error: 0.00
Seq primer: mob.RGA+ET.

Source

```
/organism="Homo sapiens"
/db_xref="GDB:1272285"
```

```
BASE COUNT      123 a      87 c      69 g      114 t  

Origin  

/db_xref="taxon:9606"  

/clone="IMAGE:346910"  

/clone_lib="Soares_fetal_heart_NBhl19W"  

/sec="unknown"  

/dev_stage="19 weeks"  

/lab_host="DH10B (ampicillin resistant)"  

/note="Organ: heart; Vector: pT73D (Pharmacia) with a  

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st  

strand cDNA was primed with a Not I - oligo(dt) primer [5'  

TGTTTCACATCTGAAGTGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3']  

double-stranded cDNA was size selected, ligated to Eco RI  

adapters (Pharmacia), digested with Not I and cloned into  

the Not I and Eco RI sites of a modified pT73 vector  

(Pharmacia). Library went through one round of  

normalization to a Cot = 5. Library constructed by  

M.Fatima Bonaldo. This library was constructed from the  

same fetus as the fetal lung library, Soares fetal lung  

NBHL19W."
```

```

alignment_scores:
  Quality: 8.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 8
  Gaps: 0
  Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x W78167/rev ..

Align seg 1/1 to reverse of: W78167 from: 1 to: 393

220 ArgGlnIlePheSerAspTyrGln 227
|||||
69 AGACAATCTTTCTGATTATCA 46

seq_name: gb_gss:AQ114866

seq_documentation_block:
LOCUS AQ114866 395 bp DNA linear GSS 29-AUG-1998
DEFINITION CIT-HSP-2375K8.TR CIT-HSP Homo sapiens genomic clone 2375K8, DNA
sequence.
ACCESSION AQ114866 GI:3490987
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 395)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
JOURNAL
COMMENT Other GSSs: CIT-HSP-2375K8.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source location/Qualifiers
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2375K8"
/clone_1ib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 102 a 92 c 93 g 108 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ114866/rev ..

Align seg 1/1 to reverse of: AQ114866 from: 1 to: 395

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
367 GTATCTGCTTTAGCGGCATCCCA 344

seq_name: gb_gss:AQ120888

seq_documentation_block:
LOCUS AQ120888 396 bp DNA linear GSS 22-SEP-1998
DEFINITION HS-3073.B1.C01.MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3073 Col-1 Row-F, DNA sequence.
ACCESSION AQ120888 GI:3498054
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 396)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,D., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL MEDLINE
COMMENT 99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3073 Row: F Column: 1
Class: BAC ends
High quality sequence stop: 396.
Location/Qualifiers
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3073 Col-1 Row-F"
/clone_1ib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones In
E-Coli DH10B"

BASE COUNT 145 a 93 c 74 g 80 t 4 others
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ120888/rev ..

Align seg 1/1 to reverse of: AQ120888 from: 1 to: 396

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
100 GTATCTGCTTTAGCGGCATCCCA 77

seq_name: gb_est2:BF908074

seq_documentation_block:
LOCUS BF908074 398 bp mRNA linear EST 18-JAN-2001
DEFINITION RC1-UT0083-091000-013-f01 UT0083 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF908074
VERSION BF908074.1 GI:12299532
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 398)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,

TITLE
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

JOURNAL
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

MEDLINE
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1-UT0083-091000-013-f01&t3=2000-10-09&t4=1)

Seq primer: puc 18 forward
High quality sequence stop: 397.

FEATURES
Location/Qualifiers
1..398
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0083"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
93 a 93 c 70 g 142 t

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BF908074/rev ..

Align seg 1/1 to reverse of: BF908074 from: 1 to: 398

61 SerTherSerLeuSerLeuArgSer 68
|||||
31 TCCACCTCCCTCAGCTGAGATCA 8

seq_name: gb_est1:AM587771

seq_documentation_block:
LOCUS AM587771 400 bp mRNA linear EST 22-MAR-2000
DEFINITION ST66F07 Pine Triplex shoot tip library Pinus taeda cDNA clone
ACCESSION AM587771
VERSION AM587771.1 GI:7274798
KEYWORDS EST
SOURCE 10b101ly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE
1 (bases 1 to 400)
Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
The Pine Gene Discovery Project
Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh

TITLE
, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@uncsu.edu

JOURNAL
Seq primer: 5' lambda Triplex2 Sequencing primer.

MEDLINE
Location/Qualifiers
1..400
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST66F07"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1: SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from apex) were collected during the spring, frozen and used for mRNA isolation. The SMART-PCR method (Clontech) was used to prepare a library from 1 ug total RNA, using the Lambda Triplex vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end."

BASE COUNT
81 a 98 c 110 g 97 t 14 others

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AM587771 ..

Align seg 1/1 to: AM587771 from: 1 to: 400

14 AspGluIlelySArgSerGlyGly 21
|||||
244 GATGAGATCAAGCGGCGGCGG 267

seq_name: gb_est1:AM902939

seq_documentation_block:
LOCUS AM902939 403 bp mRNA linear EST 24-MAY-2000
DEFINITION QV3-NN1025-100500-183-b01 NN1025 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM902939
VERSION AM902939.1 GI:8067144
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 (bases 1 to 403)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M.J., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2-QV3-NN1025-100500-183-b01&t3=2000-05-10&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 403.
Location/Qualifiers

FEATURES

source

1. 403
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NN1025"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 138 a 95 c 90 g 80 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AW902939/rev ..

Align seg 1/1 to reverse of: AW902939 from: 1 to: 403

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
74 GTCTCTCCTTAGGGGCGCATCCCA 51

seq_name: gb_est1:A1350210

seq_documentation_block:

LOCUS A1350210 408 bp mRNA linear EST 01-FEB-1999
DEFINITION G038A10.X1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1910298 3',
mRNA sequence.

ACCESSION A1350210.1 GI:4087416

VERSION EST.

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 408)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

REFERENCE NCI-CGAP

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.lnl.gov/bbrp/image/image.html

Insert Length: 566 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 379.

Location/Qualifiers

1. 408

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1lb="NCI-CGAP_Lu5"

/tissue_type="carcinoid"

/lab_host="DH10B"

BASE COUNT

/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 163 a 62 c 61 g 122 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A1350210 ..

Align seg 1/1 to: A1350210 from: 1 to: 408

63 SerLeuSerLeuArgSerAlaHis 70
|||||
227 TCATTATCCTTAGCAGATCGCAC 250

seq_name: gb_est2:BF777429

seq_documentation_block:

LOCUS BF777429 410 bp mRNA linear EST 12-JAN-2001
DEFINITION NXSL_071_E05_F NXSL (NsF xylem side wood Inclined) Pinus taeda cDNA
clone NXSL_071_E05 5', mRNA sequence.

ACCESSION BF777429 GI:12125329

VERSION EST.

KEYWORDS

SOURCE Pinus taeda

ORGANISM Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 410)

REFERENCE Sederoff, R.

AUTHORS Molecular Basis of Wood Formation in the Pine Megagenome

JOURNAL Unpublished (2000)

COMMENT Contact: Johnson, Arthur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: T3

Location/Qualifiers

1. 410

/organism="Pinus taeda"

/strain="Coastal plain loblolly pine from North Carolina"

/db_xref="taxon:3352"

/clone_1lb="NXSL_071_E05"

/clone_1lb="NXSL (NsF xylem side wood Inclined)"

/tissue_type="xylem"

/cell_type="side"

/dev_stage="juvenile"

/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI

; The library is from early (spring) wood, taken from

three six-year old trees (three different genotypes), in

the juvenile phase. These trees were induced to form side

wood by bending to a 45 degree angle and tying them to the

ground. Differentiating xylem was harvested from the sides

of the inclined stems, and a mixture of all three

genotypes was used for the library. oligo-dT primed cDNA

was directionally cloned into the EcoRI-XhoI Bluescript SK

vector arms. NOTE: The sequences contain a 'cDNA adapter'

between the EcoRI site and the start of the EST. The

adapter sequence is 'AATTCGCGACGAG'."

BASE COUNT 81 a 102 c 107 g 105 t 15 others

BASE COUNT 104 a 90 c 74 g 149 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AZ768672 ..

Align seg 1/1 to: AZ768672 from: 1 to: 417

59 TTYVAlSerThrSerLeuSerLeu 66

|||||

393 TACGTAACCACTCTCTCTCTC 416

OM of: US-09-528-682-1 to: Pending_Patents_NA_Main:* out_format : pfs
Date: Jun 18, 2002 11:51 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame-d2n.model -DEV=xlp
-O=/cgn2_1/USPFO.spool/US09528682/runat_18062002.082606_8857/app-query.fasta.1.299
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=olip2n.rnp
-GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPEL=0.000
-LOOEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=60.000
-XGAPEXT=60.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000
-YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -STAR=1
-MATRIX=oligo -TRANS=human40.cdi -LIST=1000 -DOCLIN=200
-THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs
-NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09528682.ecgn1.1.7080 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPRX -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-1
Query length: 240
Database: Pending Patents_NA_Main:*
Database sequences: 21979536
Database length: 2067452561
Search time (sec): 3375.870000

MARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
MARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

Score list:

Sequence	Strid	Orig	ZScore	EScore	Len	Documentation
/cgn2_6/pdata/2/pna/US097B.COMB.seq:US-09-724-315-4			208.00	3949.14	1.6e-210	5500
/cgn2_6/pdata/2/pna/US097B.COMB.seq:US-09-724-315-6			185.00	3527.61	4.9e-187	5488
/cgn2_6/pdata/2/pna/US097B.COMB.seq:US-09-724-315-7			184.00	3503.25	1.1e-185	782
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-3			184.00	3503.25	1.1e-185	782
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-4			111.00	2105.05	8.4e-108	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-5			86.00	1626.14	4.0e-81	711
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-6			86.00	1626.14	4.0e-81	711
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-7			26.00	477.90	3.6e-17	582
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-8			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-9			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-10			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-11			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-12			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-13			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-14			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-15			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-16			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-17			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-18			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-19			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-20			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-21			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-22			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-23			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-24			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-25			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-26			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-27			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-28			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-29			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-30			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-31			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-32			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-33			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-34			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-35			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-36			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-37			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-38			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-39			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-40			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-41			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-42			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-43			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-44			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-45			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-46			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-47			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-48			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-49			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-50			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-51			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-52			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-53			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-54			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-55			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-56			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-57			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-58			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-59			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-60			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-61			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-62			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-63			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-64			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-65			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-66			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-67			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-68			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-69			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-70			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-71			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-72			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-73			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-74			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-75			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-76			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-77			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-78			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-79			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-80			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-81			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-82			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-83			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-84			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-85			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-86			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-87			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-88			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-89			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-90			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-91			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-92			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-93			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-94			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-95			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-96			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-97			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-98			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-99			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-100			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-101			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-102			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-103			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-104			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-105			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-106			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-107			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-108			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-109			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-110			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-111			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-112			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-113			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-114			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-115			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-116			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-117			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-118			26.00	476.35	4.4e-17	723
/cgn2_6/pdata/2/pna/US099B.COMB.seq:US-09-470-124-						

[illegible]

/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-456-16766 +	8.00	136.21	389.17	382	/cgn2_6/p/ptodata/2/pna/us6017_COMB.seq:US-60-172-360-11447 -	8.00	134.61	477.74	382
/cgn2_6/p/ptodata/2/pna/us097B_COMB.seq:US-09-724-866-17756 +	8.00	136.17	391.10	37	/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-456-13106 +	8.00	134.66	478.67	37
/cgn2_6/p/ptodata/2/pna/us097B_COMB.seq:US-09-724-866-17756 +	8.00	136.17	391.10	37	/cgn2_6/p/ptodata/2/pna/us094_COMB.seq:US-09-411-999-21360 +	8.00	134.56	480.57	37
/cgn2_6/p/ptodata/2/pna/us6017_COMB.seq:US-60-171-431-10136 +	8.00	136.17	391.10	37	/cgn2_6/p/ptodata/2/pna/us0918_COMB.seq:US-09-188-316-17843 +	8.00	134.56	482.46	37
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-27350 +	8.00	136.13	393.02	3	/cgn2_6/p/ptodata/2/pna/us093_COMB.seq:US-09-332-169-18942 +	8.00	134.47	486.23	3
/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-617-33319 +	8.00	136.11	393.98	3	/cgn2_6/p/ptodata/2/pna/us097B_COMB.seq:US-09-724-750-22917 +	8.00	134.44	488.11	3
/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-617-33319 +	8.00	136.11	393.98	3	/cgn2_6/p/ptodata/2/pna/us6017_COMB.seq:US-60-171-431-23917 +	8.00	134.44	488.11	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-19645 +	8.00	136.11	393.98	3	/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26530 +	8.00	134.44	488.11	3
/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-702-12454 +	8.00	136.08	395.90	3	/cgn2_6/p/ptodata/2/pna/us093_COMB.seq:US-09-337-424-1760 -	8.00	134.40	490.94	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-27106 +	8.00	136.08	395.90	3	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-701-10587 +	8.00	134.40	490.94	3
/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-701-19421 +	8.00	136.06	396.86	3	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-701-12464 +	8.00	134.34	494.71	3
/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-701-15721 +	8.00	136.06	396.86	3	/cgn2_6/p/ptodata/2/pna/us0958A_COMB.seq:US-09-819-0914-28789 +	8.00	134.34	494.71	3
/cgn2_6/p/ptodata/2/pna/us097B_COMB.seq:US-09-724-750-101210 +	8.00	136.04	397.82	3	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.29	497.53	3
/cgn2_6/p/ptodata/2/pna/us6017_COMB.seq:US-60-171-431-10210 +	8.00	136.04	397.82	3	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-9616 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us100_COMB.seq:US-10-074-095-606 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us100_COMB.seq:US-10-074-095-606 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/cgn2_6/p/ptodata/2/pna/us6025_COMB.seq:US-60-253-457-26276 +	8.00	136.02	398.78	382	/cgn2_6/p/ptodata/2/pna/us096A_COMB.seq:US-09-605-698-6800 +	8.00	134.25	500.35	3
/									

cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-325-448-3621 +	8.00	132.98	589.01	58	/cgcn2.6/prodata/2/pna/US096.COMB	seq:US-09-660-270-1557 -	8.00	121.17	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-378-33274 +	8.00	132.95	590.86	5	/cgcn2.6/prodata/2/pna/US096A.COMB	seq:US-09-614-150-2225 -	8.00	121.01	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22797 +	8.00	132.94	591.79	5	/cgcn2.6/prodata/2/pna/US096A.COMB	seq:US-09-614-150-2225 -	8.00	120.99	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-378-12313 +	8.00	132.94	593.64	5	/cgcn2.6/prodata/2/pna/US06016.COMB	seq:US-60-157-417-1286 -	8.00	120.99	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22733 +	8.00	132.91	594.57	5	/cgcn2.6/prodata/2/pna/US06016.COMB	seq:US-60-157-417-1286 -	8.00	120.99	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-09-873-457-22834 +	8.00	132.87	597.35	5	/cgcn2.6/prodata/2/pna/US06016.COMB	seq:US-60-157-417-1286 -	8.00	120.99	2.7e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-456-32841 +	8.00	132.87	597.35	5	/cgcn2.6/prodata/2/pna/US06013.COMB	seq:US-60-138-103-7857 -	8.00	120.34	3.0e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22808 +	8.00	132.87	597.35	5	/cgcn2.6/prodata/2/pna/US06013.COMB	seq:US-60-138-103-7857 -	8.00	120.34	3.0e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-09-873-457-22834 +	8.00	132.84	599.20	5	/cgcn2.6/prodata/2/pna/US06017.COMB	seq:US-60-171-625-512 -	8.00	119.55	3.3e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22733 +	8.00	132.84	600.12	5	/cgcn2.6/prodata/2/pna/US06017.COMB	seq:US-60-171-625-512 -	8.00	119.55	3.3e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-197-872-44332 +	8.00	132.71	609.37	6	/cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-60-191-637-9973 -	8.00	119.55	3.3e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22733 +	8.00	132.69	611.22	6	/cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-60-191-637-9973 -	8.00	119.55	3.3e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-60-253-457-22733 +	8.00	132.65	613.99	6	/cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-60-191-637-9973 -	8.00	119.55	3.3e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-09-620-338-2499 +	8.00	132.65	616.76	6	/cgcn2.6/prodata/2/pna/US0933.COMB	seq:US-09-327-504-3 -	8.00	119.31	3.4e+03	39
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-10-020-338-2499 +	8.00	132.61	617.68	6	/cgcn2.6/prodata/2/pna/US096A.COMB	seq:US-09-620-338-2499 +	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-10-020-338-2499 +	8.00	132.52	625.06	6	/cgcn2.6/prodata/2/pna/US096A.COMB	seq:US-09-620-338-2499 +	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06025.COMB	seq:US-09-874-708A-5098 -	8.00	132.52	625.06	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/2/pna/US06021.COMB	seq:US-09-605-701-10353 +	8.00	132.49	626.90	6	/cgcn2.6/prodata/2/pna/US098A.COMB	seq:US-09-702-134-77597 -	8.00	119.18	3.5e+03	
cgcn2.6/prodata/											

[illegible]

[illegible]

[illegible][illegible]

```
alignment_block:
US-09-528-682-1 x US-09-724-315-4 . . .
```

Align seg 1/1 to: US-09-724-315-4 from: 1 to: 5500

5 LeuTyrArgAlaAspSerSerProProAspGluIleIysArgSerGly 21
|||||
3987 TTrATrCGGTGCTGACTCTrGACCCCAATGAAATAAAACCTCCGGAGG 4036
21 yLeuMetProArgIleHisAsnGluTyrPheAspArgIleThrGlnMetA 38
|||||
4037 TCTTATAGCCAGAGGCAATATAGTACTCTCGATrGAGAGAACCTCAATGA 4086
38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
4087 AATATrATCTTTrATATATACACGGAGAGGAAACAAACCGCTTGTGTAGA 4136
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisIle 71
|||||
4137 TATGATrAGACGGATrATGTTTCCACTCTTCTTGTAGTTGGAAATGCTCACTT 4186
71 uAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrIleTyrValI 88
|||||
4187 AGCAGACACAGTCTATATATATACAGATrATTCACACTATATATATGTGA 4236
88 lAlaIleThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyr 104
|||||
4237 TrAGCGCACACACCAATrATGTTrATGTTrATGTATrGTATrGTAGCGGTATAC 4286
105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyr 121
|||||
4287 AGCCCTCACCCATrATGAAACAGAGAGGTTTCTCGCGTTAGTGGAAATrACCATA 4336
121 rSerGlnIleTyrGlyTyrTyrArgValAsnProGluValIleAspGluA 138

4337 TTCTCGATATATGAGATGTAIGCTGTAATTTTGGTGTATTCATGCAC 4386
138 rglneuhlsarqasnaarfglutyarqasparqlytyrarqasneleuana 154
4387 GATTACATCTGACAGGAGATATAGAACCGGATATACAGAAATCTGAAT 4436
155 llealarproalagluasprglutyrarqleuadaglyherproarphn 171
4437 ATAGCTCCGGCAAGAGATGGTGTACAGATTAGACAGGTTTCCACCGGANCA 4486
171 scinalatprargglugluarproarphneishlsialarproinglucysg 188
4487 CCAAGCTTGAGAGAAAGAACCCCTGGATTCATCATGCACCAACAAAGTTGTG 4536
188 lvasnserserarqthrllethrlglyaspthrcysaengluuthrln 204
4537 GAATTCATCAACAAACAAATATACAGGTATATCTGTATATGAGAGACCCAG 4586
205 Asnleuserthrilletyrleuana 212
4587 AATCTGAGCAATATATCTCAGG 4610

seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-315-6

seq_documentation_block:

```

: Sequence 6, Application US/09724315
: GENERAL INFORMATION:
: APPLICANT: HAYNES, Joel R.
: APPLICANT: ARRINGTON, Joshua
: TITLE OF INVENTION: NOCIEIC ACID ADJUVANTS
: FILE REFERENCE: AP41
: CURRENT APPLICATION NUMBER: US/09/724,315
: CURRENT FILING DATE: 2000-11-27
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 5488
: TYPE: DNA
: ORGANISM: plasmid pcv2007
: US-09-724-315-6

```

```
alignment_scores:
  quality: 186.00      Length: 186
  ratio: 1.000        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.0000
```

```
alignment_block:
US-09-528-682-1 x US-09-724-315-6 .
```

Align seg 1/1 to: US-09-724-315-6 from: 1 to: 5488

27	HisA	G	L	T	T	P	h	e	a	s	p	r	g	L	y	t	r	h	i	m	e	s	t	a	n	i	e	a	s	L	e	t	T	r	a	s	43	
4053	C	A	A	A	T	G	A	C	T	T	C	A	T	G	A	G	A	C	T	C	A	A	G	A	A	T	T	A	A	C	T	T	T	A	G	4102		
43	p	h	i	s	A	a	r	g	L	y	t	r	h	i	m	e	s	t	a	n	i	e	a	r	g	T	r	a	s	p	a	s	p	L	e	t	r	43
4103	T	C	A	C	G	G	A	G	A	C	A	A	C	G	G	C	T	T	T	T	G	C	A	T	A	T	G	A	T	G	A	C	G	A	T	4152		
60	a	l	S	e	r	t	r	S	e	r	L	e	u	S	e	r	S	e	r	A	l	i	S	e	u	A	G	L	y	S	e	r	I	I	e	76		
4153	T	T	T	C	C	A	C	T	T	C	T	T	A	G	A	A	G	G	C	A	C	T	T	A	G	C	A	G	A	G	A	C	T	A	4202			
77	L	e	u	S	e	r	G	L	y	S	e	r	h	i	r	T	r	L	i	e	r	V	a	L	i	e	a	r	h	a	L	a	P	r	a	s	93	
4203	T	T	A	T	C	A	G	A	T	A	T	C	C	A	T	A	T	A	T	A	T	G	T	T	A	G	A	G	A	C	A	A	4252					
93	m	e	t	p	h	e	a	s	a	l	a	s	p	a	s	p	a	L	e	u	G	L	y	a	L	y	S	e	r	P	r	o	h	S	p	r	o	110
1253	T	A	G	T	T	T	T	A	A	T	A	G	A	G	T	T	T	A	G	G	C	T	A	A	C	G	C	C	T	C	A	C	C	A	4302			


```

110 lucIngIuValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly 126
|||||
4303 AACAGAGGTTTCTGCTTAGCTGAGTACCATATTCTCAGATATATGGA 4352
127 TyrPTrArgValAsnPhgGlyValIleAspGluArgLeuHisArgAsnAr 143
|||||
4353 TGGTATCGTGTATATTGGTGTGATGTAGTAAGATTACATCGTATACAG 4402
143 gGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGlu 160
|||||
4403 GGAATATAGAGACCGGTATTACAGAAATCTGAATATAGCTCCGACAGG 4452
160 spGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTTrpArgGlu 176
|||||
4453 ATGGTTACATATAGCAGGTTTCCACCGCATCACCAAGCTTGAGAGAA 4502
177 GluProTTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgTrh 193
|||||
4503 GAACCCGTGATCATCATGCACCAAGAGTTGTGAAATTCATCAAGAAC 4552
193 rIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleT 210
|||||
4553 AATTACAGGTACTTGTATATAGAGAGACCAAGATCTGAGCACATATAT 4602
210 yrlLeuArg 212
|||||
4603 ATCTCAGG 4610

seq_name: /cgn2_6/ptodata/2/pna/PCtUS.COMB.seq: PCT-US99-30747-3

seq_documentation_block:
; Sequence 3, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-3

alignment_scores:
Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-3 ..
Align seg 1/1 to: PCT-US99-30747-3 from: 1 to: 782

5 LeuTyrArgAlaAspSerArgProProAspGluIleLysArgSerGlyGly 21
|||||
72 CTCATATAGGCTGACTCTAGGCCCCCGATGATCAAGAGGTCTGAGG 121
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
|||||
122 TTCATATGCCAAGGAGACCAATAGTACTTGTATAGGGAACCAATATGA 171

```

```

38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
172 ACATCAACCTCTATGACCATGCTTAGGGGAACCCAACTGCTTTGTAGG 221
55 TyrAspArgGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
|||||
222 TATGATGATGGAATATGTTGCCACCTCCCTTAGCTTGAGAGTCTGCTACTT 271
71 uAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrVal 88
|||||
272 GCGCAGGCAATCCATCCCTCAGATATCCACCTCTACTATCATATGTGA 321
88 leaIatTrpAlaProAsnMetPheAsnValAsnAspValIleuGlyValTyr 104
|||||
322 TTGCTTACAGACCAACATATGTTCAATGTGAATGATGTGTGGAGTAT 371
105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyIleProTyr 121
|||||
372 AGCCCTCACCCATATGAGCAAGAGGTGTCTGCTTTGGGTGGAATCCCAT 421
121 rSerGlnIleTyrGlyTyrPTrArgValAsnPhgGlyValIleAspGlu 138
|||||
422 CTCCCAATCTATGATGATGTATAGGTTGAACCTTGTGTGATGATGAGA 471
138 rGleuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
472 GCGTCCATAGGAATAGGAGATATAGGACAGGTACTATAGAACCTCTAAC 521
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
|||||
522 ATACCTCCAGCAGAGAGATGTTATAGTTGGCAGGTTTCCACACAGACA 571
171 sGlnAlaTTrpArgGluGluProTTrpIleHisAlaProGlnGlyCysG 188
|||||
572 CCAAGCCTGAGAGAGAGAGCCCTGATCCACCATGACACCAAGTTGTG 621
188 ly 188
||
622 GA 623

seq_name: /cgn2_6/ptodata/2/pna/US094.COMB.seq: US-09-470-124-3

seq_documentation_block:
; Sequence 3, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-3

alignment_scores:
Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x US-09-470-124-3 ..
Align seg 1/1 to: US-09-470-124-3 from: 1 to: 782

```

5 LeuTyraArgAlaaspSerArgProProaspGluLeuLysArgSerGly 21
|||||
72 CTCTATAGGCGCTGACTAGAGCCCGCCAGATGATCAAGAGCTGTGAGG 121
21 yLeuMetProArgGlyHisAsnGlyuTyrrPheAspArgGlyThrLinneta 38
|||||
122 TCTCATGCCAAGGGGACACAATGAGTACTTGTATGAGGGAAACCAATATGA 171
38 snIleAsnLeuTyraSPHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
172 ACATCAACCTCTATGACCATGCTAGGGGAACCAACTGGCTTTGTGAGG 221
55 TyraSPaspGlyTyrrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
|||||
222 TATGATGATGATATGTGTCACCTCCCTTAGCTTGAGGTCGCTCACTT 271
71 uAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrrIleTyrrVal 88
|||||
272 GGCAGGACAATCCATCCCTCAGGATACTCCACTACTACATCTATGTGA 321
88 leaIaThrAlaProAsnMetPheAsnValAsnAspValIleuGlyValTyrr 104
|||||
322 TTGCTACAGCACAACATGTTCAATGTGATGTGTGGAGGTAT 371
105 SerProHisProTyrrGluGlnGluValSerAlaLeuGlyIleProTy 121
|||||
372 AGCCCTCACCCTATGAGCAAGAGGTGTGCTGTTGGTGGAATCCATA 421
121 rSerGlnIleTyrrGlyTyrrArgValAsnPheGlyValIleAspGlu 138
|||||
422 CTCCCAATCTATGATGATGATGAGGTGAACTTGTGTGATGATGA 471
138 rGleuHisArgAsnArgIleuTyrrArgAspArgTyrrTyrrArgAsnLeuAsn 154
|||||
472 GGTCCATAGGATAGGAGATAGGAGGACAGTACTATAGAACCTCAAC 521
155 IleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPheProProAspH 171
|||||
522 ATACTCCAGCAGAGAGATGTTATAGTTGGCAGTTTCCACACAGACA 571
171 sGlaIaTrpArgGluGluProTrpIleHisAlaProGlnGlyCysG 188
|||||
572 CCAAGCCTGGAGGAGAGACCCCTGATCCACATGACACACAGAGTTGTG 621
188 Ly 188
||
622 GA 623

```

seq_name: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-950-335A-5

```

seq_documentation_block:
; Sequence 5, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTION
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-950-335A-5

```

alignment_scores:
Quality: 111.00 Length: 111
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-950-335A-5 ..
Align seg 1/1 to: US-09-950-335A-5 from: 1 to: 723

```

1 AsnGlyAspArgLeuTyraArgAlaaspSerArgProProaspGluLeu 17
|||||
1 AATGCCACAGATTTATCCCTGCTGACTTAGACCCCGCCAGATGAATATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyuTyrrPheAspArg 34
|||||
51 ACGTTCGGAGGCTTATGCCAGAGGGCATATGAGTACTTGATGATAGAG 100
34 yThrGlnMetAsnIleAsnLeuTyraSPHisAlaArgGlyThrGlnThr 50
|||||
101 GAACCTCAATGATATTTATCTTATGATGATCAGCGAGAGACACAAC 150
51 GlyPheValArgTyraSPaspGlyTyrrValSerThrSerLeuSerLeu 67
|||||
151 GCGTTTGTCCAGATATGATGAGGATATGTTCCACTTCTCTAGTTTGA 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrr 84
|||||
201 AAGTCTCACTTAGCAGCAGACGCTATATATCAAGATATCCACTTACT 250
84 yIleTyrrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
251 ATATATATGTTATAGCAGACAGACCAATATGTTATATGTTATATGATGA 300
101 LeuGlyValTyrrSerProHisProTyrrGluGln 111
|||||
301 TTAGCGCTATACAGCCCTCACCCTATATGAAACAG 333

```

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-5

```

seq_documentation_block:
; Sequence 5, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera toxin and of the toxin Lt, Their Preparation and
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) (655-3542

```



```
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..711
US-08-256-003-5

alignment_scores:
    Quality: 86.00      Length: 86
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-256-003-5 ..

Align seg 1/1 to: US-08-256-003-5 from: 1 to: 711

103 ValTYSerProHISProTYrGluGluGluValSerAlaLeuGlyIyl 119
|||||
295 GTATACAGCCCTCACCATATGACAGAGAGGTTTCGCGTTAGTGGAAT 344
119 eProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhcGlyValIleA 136
|||||
345 ACCATATTCAGATATATGATGATGATGATGATGATGATGATGATGATG 394
136 sPGluArGLeuHISArGAsnArGLeuTYrArgAspArgTYrTYrArgAsn 152
|||||
395 ATGACGATTACATCTGTACAGAGGAAATATAGAGCCGTTATTACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPheProP 169
|||||
445 CTGATATATAGCTCCGCGACAGAGATGTTACAGATTGACAGTTCCACAC 494
169 oAspHISGlnAlaTPArgGluGluProTPriIeHISAlaProGlnG 186
|||||
495 GGATCACCAAGCTTGAGAGAGAAAGAACCTGATTCATCATGCACCAAG 544
186 lYcysGly 188
|||||
545 GTTGTGGA 552

seq_name: /cgn2/6/plodata/2/pna/US090_COMB.seq:US-09-044-696-1

seq_documentation_block:
; Sequence 1, Application US/09044696
; GENERAL INFORMATION:
; APPLICANT: BARCHFIELD, GAIL
; APPLICANT: DEL GUIDICE, GIUSEPPE
; APPLICANT: RAPPUOLI, RINO
; TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
; TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
; ADDRESSSEE: R440
; STREET: P.O. BOX 8097
; CITY: EMERYVILLE
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/044,696
```

```
; FILING DATE: 18-MAR-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,227
; FILING DATE: 21-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 1393.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 655-8730
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
US-09-044-696-1

alignment_scores:
    Quality: 86.00      Length: 86
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-044-696-1 ..

Align seg 1/1 to: US-09-044-696-1 from: 1 to: 711

103 ValTYSerProHISProTYrGluGluGluValSerAlaLeuGlyIyl 119
|||||
295 GTATACAGCCCTCACCATATGACAGAGAGGTTTCGCGTTAGTGGAAT 344
119 eProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhcGlyValIleA 136
|||||
345 ACCATATTCAGATATATGATGATGATGATGATGATGATGATGATGATG 394
136 sPGluArGLeuHISArGAsnArGLeuTYrArgAspArgTYrTYrArgAsn 152
|||||
395 ATGACGATTACATCTGTACAGAGGAAATATAGAGCCGTTATTACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPheProP 169
|||||
445 CTGATATATAGCTCCGCGACAGAGATGTTACAGATTGACAGTTCCACAC 494
169 oAspHISGlnAlaTPArgGluGluProTPriIeHISAlaProGlnG 186
|||||
495 GGATCACCAAGCTTGAGAGAGAAAGAACCTGATTCATCATGCACCAAG 544
186 lYcysGly 188
|||||
545 GTTGTGGA 552

seq_name: /cgn2/6/plodata/2/pna/US099B_COMB.seq:US-09-950-335A-2

seq_documentation_block:
; Sequence 2, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
```


LENGTH: 582
TYPE: DNA
ORGANISM: Vibrio cholerae
US-09-950-335A-2

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-950-335A-2 ..

Align seg 1/1 to: US-09-950-335A-2 from: 1 to: 582

29 GUTYRPHASPARGLYTHRGINMETASNILEASNLEUTYRASPISAL 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGAAATATCAACCTTATGATCATGC 134
45 AARGGLYTHRGINTHRGLYPHEVALARG 54
|||||
135 AAGAGSACTCAGACGCGATTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-7

seq_documentation_block:

Sequence 7, Application US/08256003
GENERAL INFORMATION:
APPLICANT: Domenighini, Mario
APPLICANT: Rappuoli, Rino
TITLE OF INVENTION: Immuno-genic Detoxified Mutants of
TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,003
FILING DATE: 11-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..723
US-08-256-003-7

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-256-003-7 ..

Align seg 1/1 to: US-08-256-003-7 from: 1 to: 723

29 GUTYRPHASPARGLYTHRGINMETASNILEASNLEUTYRASPISAL 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGAAATATCAACCTTATGATCATGC 134
45 AARGGLYTHRGINTHRGLYPHEVALARG 54
|||||
135 AAGAGSACTCAGACGCGATTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-044-696-3

seq_documentation_block:

Sequence 3, Application US/09044696
GENERAL INFORMATION:
APPLICANT: BARCHFIELD, GAIL
APPLICANT: DEL GIUDICE, GIUSEPPE
TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
STREET: P.O. BOX 8097
CITY: EMERYVILLE
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/044,696
FILING DATE: 18-MAR-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,227
FILING DATE: 21-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 1393,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 655-8730
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..720
NAME/KEY: CDS
LOCATION: 1..720
US-09-044-696-3

alignment_scores: Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-044-696-3 ..

Align seg 1/1 to: US-09-044-696-3 from: 1 to: 723

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATATGATCATGCG 134
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
135 AAGGGAACCTCAGACGGGATTGTGTAGG 162
```

seq_name: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-950-335A-1

seq_documentation_block:

```
; Sequence 1, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTION
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-09-950-335A-1
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-950-335A-1 ..

Align seg 1/1 to: US-09-950-335A-1 from: 1 to: 723

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATATGATCATGCG 134
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
135 AAGGGAACCTCAGACGGGATTGTGTAGG 162
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-45

seq_documentation_block:

```
; Sequence 45, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: (1)..(777)

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-45
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-45 ..

Align seg 1/1 to: PCT-US99-30747-45 from: 1 to: 777

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACCGAGGTACTCAGATGACATCAACCTTATATGACCATGCG 188
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAAACTGATTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-47

seq_documentation_block:

```
; Sequence 47, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-47
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-47 ..

Align seg 1/1 to: PCT-US99-30747-47 from: 1 to: 777

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACCGAGGTACTCAGATGACATCAACCTTATATGACCATGCG 188
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAAACTGATTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-49


```
seq_documentation_block:
; Sequence 49, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Transgenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-49
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:
US-09-528-682-1 x PCT-US99-30747-49 ..

Align seg 1/1 to: PCT-US99-30747-49 from: 1 to: 777

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGGTACTCGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAGACTCAACGATGATTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-51

seq_documentation_block:

```
; Sequence 51, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-51
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x PCT-US99-30747-51 ..

Align seg 1/1 to: PCT-US99-30747-51 from: 1 to: 777

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGGTACTCGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAGACTCAACGATGATTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-07-694-733-1

seq_documentation_block:

```
; Sequence 1, Application US/07694733
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: Recombinant DNA-derived
; TITLE OF INVENTION: Cholera Toxin
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1840 Denavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh OS 7.0.
; SOFTWARE: Microsoft Word Version 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/694,733
; FILING DATE: 19910502
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double stranded
; TOPOLOGY: circular
; US-07-694-733-1
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:
US-09-528-682-1 x US-07-694-733-1 ..

Align seg 1/1 to: US-07-694-733-1 from: 1 to: 777

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGGTACTCGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAGACTCAACGATGATTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-435-605-1

```
seq_documentation_block:
; Sequence 1, Application US/08435605
```



```

; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: Recombinant DNA-derived
; TITLE OF INVENTION: Cholera Toxin
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh OS 7.0.
; SOFTWARE: Microsoft Word Version 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/694,733
; FILING DATE: 02-MAY-1991
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: circular
; US-08-435-605-1

```

```

alignment_scores:
      Quality: 26.00      Length: 26
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

US-09-528-682-1 x US-08-435-605-1 ..

Align seg 1/1 to: US-08-435-605-1 from: 1 to: 777

```

29 GltutyrPheasparglythrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAATATCAACCTTTATGTCATGC 188
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACGAGGATTGTGAGG 216

```

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-45

```

seq_documentation_block:
; Sequence 45, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arntzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS

```

```

; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
; US-09-470-124-45

```

```

alignment_scores:
      Quality: 26.00      Length: 26
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

US-09-528-682-1 x US-09-470-124-45 ..

Align seg 1/1 to: US-09-470-124-45 from: 1 to: 777

```

29 GltutyrPheasparglythrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAATATCAACCTTTATGTCATGC 188
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACGAGGATTGTGAGG 216

```

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-47

seq_documentation_block:

Sequence 47, Application US/09470124

GENERAL INFORMATION:

APPLICANT: Mason

APPLICANT: Arntzen

TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In

TITLE OF INVENTION: Transgenic Plants

FILE REFERENCE: 4868/84454

CURRENT APPLICATION NUMBER: US/09/470,124

CURRENT FILING DATE: 1999-12-24

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 47

LENGTH: 777

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(777)

OTHER INFORMATION: Description of Artificial Sequence:V. cholerae

OTHER INFORMATION: cholera toxin gene mutagenized to optimize

OTHER INFORMATION: expression in plants.

US-09-470-124-47

```

alignment_scores:
      Quality: 26.00      Length: 26
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

US-09-528-682-1 x US-09-470-124-47 ..

Align seg 1/1 to: US-09-470-124-47 from: 1 to: 777

```

29 GltutyrPheasparglythrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAATATCAACCTTTATGTCATGC 188
|||||
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACGAGGATTGTGAGG 216

```

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-49

seq_documentation_block:
; Sequence 49, Application US/09470124
; GENERAL INFORMATION:

; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-49

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-470-124-49 ..

Align seg 1/1 to: US-09-470-124-49 from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTCAGATGAACTCAACCTTTATGACCATGC 188
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
189 AAGGGGAACCTCAACCTGATTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:us-09-470-124-51

seq_documentation_block:

; Sequence 51, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-51

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-470-124-51 ..

Align seg 1/1 to: US-09-470-124-51 from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTCAGATGAACTCAACCTTTATGACCATGC 188
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
189 AAGGGGAACCTCAACCTGATTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US97-11719-12

seq_documentation_block:

; Sequence 12, Application PC/TUS9711719
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS
; TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/11719
; FILING DATE: 03-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/674,895
; FILING DATE: 03-JUL-1996
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1956 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PCT-US97-11719-12

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x PCT-US97-11719-12 ..

Align seg 1/1 to: PCT-US97-11719-12 from: 1 to: 1956

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
1053 GAGTACTTGTGACAGGGGTACTCAGATGAACTCAACCTTTATGATCATGC 1102
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
1103 AAGAGGAACCTCAGACGGGATTGTAGG 1130

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-08-674-895-12

seq_documentation_block:

; Sequence 12, Application US/08674895
; GENERAL INFORMATION:

APPLICANT: Blell, Jeffrey D.
APPLICANT: Bookbinder, Louis H.
APPLICANT: Hehn, Mich B.
TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS
TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 North Torrey Pines Road
CITY: La Jolla
STATE: California
COUNTRY: United States
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,895
FILING DATE: 03-JUN-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 529.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-674-895-12

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-674-895-12 ..

Align seg 1/1 to: US-08-674-895-12 from: 1 to: 1956

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
1053 GAGTACTTGTGACCGAGGTACTCAATGAATGATCAACCTTATGATCATGC 1102
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
1103 AAGAGAACTCAGACGCGGATTGTGTAGG 1130

seq_name: /cgn2_6/plodata/2/pna/PCTUS_COMB.seq:PCT-US01-08582-1

seq_documentation_block:
Sequence 1, Application PC/TUS0108582
GENERAL INFORMATION:
APPLICANT: UAB Research Foundation
TITLE OF INVENTION: Chimeric Nontoxic Mutants of
Enterotoxins as Mucosal Adjuvants for Cell-Mediated or
Humoral Immunity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hendricks and Assoc
STREET: P.O. Box 2509
CITY: Fairfax

STATE: VA
COUNTRY: US
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US01/08582
FILING DATE: 16-Mar-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenn
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: MCG-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/425-8405
TELEFAX: 703/425-8406
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
PCT-US01-08582-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US01-08582-1 ..

Align seg 1/1 to: PCT-US01-08582-1 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGTGACCGAGGTACTCAATGAATGATCAACCTTATGATCATGC 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGAACTCAGACGCGGATTGTGTAGG 733

seq_name: /cgn2_6/plodata/2/pna/PCTUS_COMB.seq:PCT-US01-08582-2

seq_documentation_block:
Sequence 2, Application PC/TUS0108582
GENERAL INFORMATION:
APPLICANT: UAB Research Foundation
TITLE OF INVENTION: Chimeric Nontoxic Mutants of
Enterotoxins as Mucosal Adjuvants for Cell-Mediated or
Humoral Immunity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hendricks and Assoc
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: US
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US01/08582
FILING DATE: 16-Mar-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: MCG-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/425-8405
TELEFAX: 703/425-8406
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
PCT-US01-08582-2

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US01-08582-2 ..

Align seq 1/1 to: PCT-US01-08582-2 from: 1 to: 2022

29 GlutyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||
656 GAGTACTTTGACGAGGTACTCAATGATATCAACCTTTATGATCATGC 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGAACTCAGACGGGATTGTAGG 733

seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-315-3

seq_documentation_block:

Sequence 3, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: APF41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 5488
TYPE: DNA
ORGANISM: plasmid pPJV2006
US-09-724-315-3

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-724-315-3 ..

Align seq 1/1 to: US-09-724-315-3 from: 1 to: 5488

29 GlutyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||

4059 GAGTACTTTGACGAGGTACTCAATGATATCAACCTTTATGATCATGC 4108
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
4109 AAGAGAACTCAGACGGGATTGTAGG 4136

seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-315-1

seq_documentation_block:

Sequence 1, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: APF41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5500
TYPE: DNA
ORGANISM: pPJV2002 plasmid
US-09-724-315-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-724-315-1 ..

Align seq 1/1 to: US-09-724-315-1 from: 1 to: 5500

29 GlutyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyraSPHSAI 45
|||||
4059 GAGTACTTTGACGAGGTACTCAATGATATCAACCTTTATGATCATGC 4108
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
4109 AAGAGAACTCAGACGGGATTGTAGG 4136

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-08-670-974-6

seq_documentation_block:

Sequence 6, Application US/08670974
GENERAL INFORMATION:
APPLICANT: Mekalanos, John J.
TITLE OF INVENTION: NOVEL CHOLERA VECTORS, VACCINES,
TITLE OF INVENTION: AND METHODS FOR ANTIGEN DELIVERY IN GRAM-NEGATIVE
TITLE OF INVENTION: BACTERIA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,974
FILING DATE: 26-JUN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00742/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 6943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
US-08-670-974-6

Alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x US-08-670-974-6 ..

Align seg 1/1 to: US-08-670-974-6 from: 1 to: 6943

29 GIUTYRPhEaSPaRGlyThrGlnMetAsnIleAsnLeuTYrAspHisAl 45
|||||
5829 GAGTACTTTGACCGAGGTACTCAATGAAATATCAACCTTATATCATGTC 5878
|||||
45 AARGlyThrGlnThcGlyPheValArg 54
|||||
5879 AAGAGGAACTCAGACGCGATTGTGTAGG 5906

seq_name: /cgn2_6/ptodata/2/pna/PCrUS_COMB.seq:PCT-US01-16904-1

seq_documentation_block:
; Sequence 1, Application PC/PTUS0116904
; GENERAL INFORMATION:
; APPLICANT: University of Maryland Biotechnology Institute
; APPLICANT: Powell, Robert J.
; APPLICANT: Hone, David M.
; TITLE OF INVENTION: Methods for Introducing and Expressing Genes in Animal Cells, and
; TITLE OF INVENTION: Bacterial Biebs for Use in Same
; FILE REFERENCE: 4115-121 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/16904
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/206,994
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 7586
; TYPE: DNA
; ORGANISM: Cholera Toxin A1 Subunit Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6184)..(6184)
; OTHER INFORMATION: n can be any one of a, c, g, and t.
; NAME/KEY: misc_feature
; LOCATION: (6226)..(7053)
; OTHER INFORMATION: n can be any one of a, c, g, and t.
; NAME/KEY: misc_feature
; LOCATION: (6224)..(6224)
; OTHER INFORMATION: n can be any one of a, c, g, and t.
; PCT-US01-16904-1

Alignment_scores:
Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x PCT-US01-16904-1 ..

Align seg 1/1 to: PCT-US01-16904-1 from: 1 to: 7586

29 GIUTYRPhEaSPaRGlyThrGlnMetAsnIleAsnLeuTYrAspHisAl 45
|||||
87 GAGTACTTTGACCGAGGTACTCAATGAAATATCAACCTTATATCATGTC 136
|||||
45 AARGlyThrGlnThcGlyPheValArg 54
|||||
137 AAGGAGAACTCAGACGCGATTGTGTAGG 164

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-040-990-1

seq_documentation_block:
; Sequence 1, Application US/09040990
; GENERAL INFORMATION:
; APPLICANT: Clements, J.
; APPLICANT: Friede, M.
; TITLE OF INVENTION: USE OF MUTANT ENTEROTOXIN WITH EXCESS B-SUBUNIT AS AN
; TITLE OF INVENTION: ADJUVANT
; FILE REFERENCE: 5113-052
; CURRENT APPLICATION NUMBER: US/09/040,990
; CURRENT FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(45)
US-09-040-990-1

Alignment_scores:
Quality: 15.00 Length: 15
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x US-09-040-990-1 ..

Align seg 1/1 to: US-09-040-990-1 from: 1 to: 45

186 GLYcysGlyAsnSerArgThrIleThcGlyAspThrCysAsn 200
|||||
1 GGTGTGGAATTCATCAAGAACATTAACAGTGATCTGTAT 25

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-365-530-1

seq_documentation_block:
; Sequence 1, Application US/09365530
; GENERAL INFORMATION:
; APPLICANT: Clements, John D.
; APPLICANT: Dickinson, Bonny L.
; TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A
; TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/365,530
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/296,848
; FILING DATE: 26-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5113-046
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-09-365-530-1

```

```

alignment_scores:
  Quality: 15.00      Length: 15
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x US-09-365-530-1 ..

```

```

Align seg 1/1 to: US-09-365-530-1 from: 1 to: 45

```

```

186 GlycylglyAsnSerArgThrIleHrcGlyAspThrCysAsn 200
|||||
1 GGTGTGCAATTCATCAAGACAATTCACAGTGCTACTTGTAAAT 45

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCUTUS_COMB.seq:PCT-US99-30747-8

```

```

seq_documentation_block:
; Sequence 8, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-8

```

```

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-8 ..

```

```

Align seg 1/1 to: PCT-US99-30747-8 from: 1 to: 40

```

```

19 SerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPhe 31
|||||
2 TCTGGAGCTCTATGCCAAGGACACAAATGAGTACTTT 40

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCUTUS_COMB.seq:PCT-US99-30747-9

```

```

seq_documentation_block:
; Sequence 9, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-9

```

```

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-9 ..

```

```

Align seg 1/1 to: PCT-US99-30747-9 from: 1 to: 40

```

```

32 AspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHis 44
|||||
1 GATGGGGGAAACCCAAATGACATCAACCTCTATGACCAT 39

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCUTUS_COMB.seq:PCT-US99-30747-11

```

```

seq_documentation_block:
; Sequence 11, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-11

```

```

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x PCT-US99-30747-11 ..

Align seg 1/1 to: PCT-US99-30747-11 from: 1 to: 40

59 TyrValSerThrSerLeuSerIleuArgSerAlaHisIleu 71
|||||
2 TATGTGTCCACCTCTGCTTACGCTGAGTCTGCCTCCTTG 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-12

seq_documentation_block:

; Sequence 12, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-12

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x PCT-US99-30747-12 ..

Align seg 1/1 to: PCT-US99-30747-12 from: 1 to: 40

72 AlaGlyGlnSerIleuSerGlyTyrSerThrTyrTyr 84
|||||
1 GCAGGACAAATCCATCCTCTCAGGATACCTCCACTACTAC 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-14

seq_documentation_block:

; Sequence 14, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-14

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x PCT-US99-30747-14 ..

Align seg 1/1 to: PCT-US99-30747-14 from: 1 to: 40

99 AspValLeuGlyValTyrSerProHisProTyrGluGln 111
|||||
2 GATGTGTGGAGAGTATAGCCCTCACCACCATATGAGCAA 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-15

seq_documentation_block:

; Sequence 15, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-15

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x PCT-US99-30747-15 ..

Align seg 1/1 to: PCT-US99-30747-15 from: 1 to: 40

112 GluValSerAlaLeuGlyGlyIleProTyrSerGlnIle 124
|||||
1 GAGGTGTCTGCTTGGGTGCATTCCTACTCCCAATTC 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-17

seq_documentation_block:

; Sequence 17, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-17


```
alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-17  ..

Align seg 1/1 to: PCT-US99-30747-17 from: 1 to: 40

139 LeuHlsArGAsnArGluTYrArGAspArGTYrTYrArg 151
|||||
2 CTCATAGAGATAGGAGATATAGGACAGAGATAGGAC 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-18

seq_documentation_block:
; Sequence 18, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 18
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-18

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-18  ..

Align seg 1/1 to: PCT-US99-30747-18 from: 1 to: 40

152 AsnLeuAsnIleAlaProAlaGluAspGluTYrArGleu 164
|||||
1 AACCTCAACATAGCTCCAGCAGAGAGATGCTATAGGTTG 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-21

seq_documentation_block:
; Sequence 21, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 21
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-21

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-21  ..

Align seg 1/1 to: PCT-US99-30747-21 from: 1 to: 40

192 ArgThrIleThrgIAspThrCysAsnGluGluThrgln 204
|||||
1 AGGACATCAAGAGTGACACTTCGATGAGAGACCCAA 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-23

seq_documentation_block:
; Sequence 23, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 23
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-23

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-23  ..

Align seg 1/1 to: PCT-US99-30747-23 from: 1 to: 40

219 LysArGlnIlePheSerAspTYrGlnSerGluValAsp 231
|||||
2 AAGAGCGCAATCTTCTCAGACTACCAATCAGAGGTGAC 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-28

seq_documentation_block:
; Sequence 28, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 28
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-28
```



```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-28

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-28/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-28 from: 1 to: 40

199 CysAnGluGluThrGlnAsnLeuSerThrIleTyrLeu 211
|||||
39 TGCAATGAGAGAGACCAGCTTAGCAGCATCTACTCT 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-30

seq_documentation_block:
; Sequence 30, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-30

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-30/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-30 from: 1 to: 40

172 GlnAlaThrparGluGluProTyrPheHisHisAlaPro 184
|||||
40 CAAGCCTGGAGGAGGCCCTGATCCACCATGCACCA 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-31

seq_documentation_block:
; Sequence 31, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 31
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-31

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-31/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-31 from: 1 to: 40

159 GluAspGlyTyrArgLeuAlaGlyPheProPronAspHis 171
|||||
39 GAGGATGTTATAGTTGCGAGGTTTCCACAGACGAC 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-33

seq_documentation_block:
; Sequence 33, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-33

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-33/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-33 from: 1 to: 40

132 PheGlyValIleAspArgLeuHisArgAsnArgGlu 144
|||||
40 TTTCGTGATTTGATGAGAGCGCTCATAGAGATAGCGAG 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-34

seq_documentation_block:
; Sequence 34, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
```



```
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
PCT-US99-30747-34

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-34/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-34 from: 1 to: 40

119 ILEProTyrSerGlnIleTyrGlyTyrTyrGValasn 131
|||||
39 ATCCCATCTCCCAATCTATGATGCTATAGCGTGAC 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-36

seq_documentation_block:
; Sequence 36, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-36

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-36/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-36 from: 1 to: 40

92 ProAsnMetPheAsnValAsnAspValLeuGlyValTyr 104
|||||
40 CCAACATGTTCAATGATGATGATGTTGGAGTGTAT 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-37

seq_documentation_block:
; Sequence 37, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
```

```
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
PCT-US99-30747-37

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-37/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-37 from: 1 to: 40

79 GLYTYSerThrTyrTyrIleTyrValIleAlaThrAla 91
|||||
39 GGATACCTCCACCTACATCATCTATGATGCTACACCA 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-39

seq_documentation_block:
; Sequence 39, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-39

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-39/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-39 from: 1 to: 40

52 PheValArgTyrAspAspGlyTyrValSerThrSerIeu 64
|||||
40 TTGTGTAGGTATGATGATGATGATGTCACCTCCCTT 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-40

seq_documentation_block:
; Sequence 40, Application PC/TUS9930747
```


; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-40

alignment_scores:
Quality: 13.00 Length: 13
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x PCT-US99-30747-40/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-40 from: 1 to: 40

39 IleasnleuTYrAspHisAlaArgGlyThrGlnThrGly 51
|||||
39 ATCACCTCTATGACCATGCTAGGGCAACCAACTGTC 1

[illegible][illegible]

[illegible]

[illegible]

[illegible]

[illegible][illegible]

/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-142-766-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-139-980-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-274-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-471-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-473-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-806-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-807-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-808-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-809-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-810-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-861-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-863-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-140-921-409 + 7.00 102.07 1.4e+03 3

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-1

seq_documentation_block:

; Sequence 1, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshifumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-809-033A-1

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-809-033A-1 ..

Align seg 1/1 to: US-09-809-033A-1 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGGACCGAGGTACTCAATGATATCATCCTTATGATCATGC 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGGAATCTCAGACGGGATTTGTAGG 733

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-2

seq_documentation_block:

; Sequence 2, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshifumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-809-033A-2

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-809-033A-2 ..

Align seg 1/1 to: US-09-809-033A-2 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGGACCGAGGTACTCAATGATATCATCCTTATGATCATGC 705

45 AAGGAGTACGAGGATTTGTTAG 54
 |||||||
 706 AAGGAGTACGAGGATTTGTTAG 733

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12399

```
seq_documentation_block:
; Sequence 12399, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12399
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12399
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-105-299-12399/rev ..

Align seg 1/1 to reverse of: US-10-105-299-12399 from: 1 to: 406

75 SerIleuSerGlyTyrSerThr 82
 |||||||
 100 TCATCCTGAGTGATACAGCACA 77

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-242566

```
seq_documentation_block:
; Sequence 242566, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242566
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-242566
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-67507/rev ..

Align seg 1/1 to reverse of: US-10-027-632-67507 from: 1 to: 651

61 SerThrSerIleuSerArgSer 68
 |||||||
 390 TCACCTCTCTGAGCCTCAGATCC 367

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203067

```
seq_documentation_block:
; Sequence 203067, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203067
```

alignment_scores:

Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-242566 ..

Align seg 1/1 to: US-10-027-632-242566 from: 1 to: 631

119 IleProTyrSerGlnIleTyrGly 126
 |||||||
 383 ATTCCTTACTCTCAGATCTATGGA 406

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-67507

```
seq_documentation_block:
; Sequence 67507, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67507
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-67507
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-67507/rev ..

Align seg 1/1 to reverse of: US-10-027-632-67507 from: 1 to: 651

61 SerThrSerIleuSerArgSer 68
 |||||||
 390 TCACCTCTCTGAGCCTCAGATCC 367

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203067

```
seq_documentation_block:
; Sequence 203067, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203067
```



```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1998-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-203067
```

```
alignment_scores:
  Quality: 8.00      Length: 8
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-203067/rev ..

Align seq 1/1 to reverse of: US-10-027-632-203067 from: 1 to: 672

```
113 ValSerAlaLeuGlyIlePro 120
|||||
645 GTATCTGCTTAGGGGCGCATCCCA 622
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-8851

```
seq_documentation_block:
; Sequence 8851, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8851
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)...('797)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-8851
```

```
alignment_scores:
  Quality: 8.00      Length: 8
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-8851/rev ..

Align seq 1/1 to reverse of: US-10-027-632-8851 from: 1 to: 797

```
68 SerAlaHisLeuAlaGlyGlnSer 75
|||||
277 AGTCCCATCTGCGTGGCGCAGAGC 254
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-8852

```
seq_documentation_block:
; Sequence 8852, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8852
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...('797)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-8852
```

```
alignment_scores:
  Quality: 8.00      Length: 8
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-8852/rev ..

Align seq 1/1 to reverse of: US-10-027-632-8852 from: 1 to: 797

```
68 SerAlaHisLeuAlaGlyGlnSer 75
|||||
277 AGTCCCATCTGCGTGGCGCAGAGC 254
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-125010


```
seq_documentation_block:
; Sequence 125010 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125010
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-125010

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-125010 ..

Align seg 1/1 to: US-10-027-632-125010 from: 1 to: 892
65 SerLeuArgSerAlaHisLeuAla 72
|||||
131 TCCTTGAGAGTCCCATCTGCC 154

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-323968

seq_documentation_block:
; Sequence 323968 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323968
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323968

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-323968 ..

Align seg 1/1 to: US-10-027-632-323968 from: 1 to: 970
65 SerLeuArgSerAlaHisLeuAla 72
|||||
221 TCCTTGAGAGTCCCATCTGCC 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258164

seq_documentation_block:
; Sequence 258164 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258164
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258164

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-258164/rev ..

Align seg 1/1 to reverse of: US-10-027-632-258164 from: 1 to: 2211
150 TyTArGAsnLeuAsnIleAlaPro 157
|||||
819 TACAGAAATCTGAATATTGCTCCG 796

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258165

seq_documentation_block:
; Sequence 258165 Application US/10027632
```

```
; SEQ ID NO 323968
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323968

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-323968 ..

Align seg 1/1 to: US-10-027-632-323968 from: 1 to: 970
65 SerLeuArgSerAlaHisLeuAla 72
|||||
221 TCCTTGAGAGTCCCATCTGCC 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258164

seq_documentation_block:
; Sequence 258164 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258164
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258164

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-258164/rev ..

Align seg 1/1 to reverse of: US-10-027-632-258164 from: 1 to: 2211
150 TyTArGAsnLeuAsnIleAlaPro 157
|||||
819 TACAGAAATCTGAATATTGCTCCG 796

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258165

seq_documentation_block:
; Sequence 258165 Application US/10027632
```



```

; TYPE: DNA
; ORGANISM: Human
US-10-027-632-112146

alignment_scores:
    Quality:      8 00
    Ratio:        1.000
    Percent Similarity: 100.000
    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-112146  ..

Align seg 1/1 to: US-10-027-632-112146 from: 1 to: 2732

162 TyArgIleuAlaIagIypHePro 169
|||||
193 TACAGGCTTGACAGGGGTCCACCT 216

seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:US-10-138-381-266

seq_documentation_block:
; Sequence 26, Application US/10138381
; GENERAL INFORMATION:
; APPLICANT: NAKAYAMA, Tomoko
; APPLICANT: TADA, Jun
; APPLICANT: FUKUSHIMA, Shigeru
; APPLICANT: OHASHI, Tetsuo

```

```

FILE REFERENCE: 1422-0430P
CURRENT APPLICATION NUMBER: US/10/138,381
PRIOR FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US/09/614,681
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 08/968,046
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 08/328,710
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: 6-48174 JAPAN
PRIOR FILING DATE: 1994-03-18
PRIOR APPLICATION NUMBER: 6-30277 JAPAN
PRIOR FILING DATE: 1994-02-28
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 26
LENGTH: 22
TYPE: DNA
ORGANISM: Vibrio cholerae
US-10-138-381-26

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment block:
US-09-528-682-1 x US-10-138-381-26/rev ..

Align seg 1/1 to reverse of: US-10-138-381-26 from: 1 to: 22

122 SerGlnIleTyrcGlyTrrTyr 128
|||||
21 TCCCAATATATGCGATGGTAT 1

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-53058
; documentation block:
; Sequence 53058, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```



```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 53058
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-53058
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-53058/rev ..

Align seg 1/1 to reverse of: US-10-027-632-53058 from: 1 to: 93

```
66 leuArgSerIaHsleuA 72
|||||
27 TTAGATCCGCGACATTAGCG 7
```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-20665

seq_documentation_block:

```
; Sequence 20665, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullaly, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO: 20665
; LENGTH: 102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00919904
US-09-539-331D-20665
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-09-539-331D-20665/rev ..

Align seg 1/1 to reverse of: US-09-539-331D-20665 from: 1 to: 102

```
63 SerIeuSerIeuArgSerA 69
|||||
29 AGCCTGAGTCTGAGAGTGC 9
```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-14267

seq_documentation_block:

```
; Sequence 14267, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO: 14267
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700908892H1
US-09-975-254-14267
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-09-975-254-14267 ..

Align seg 1/1 to: US-09-975-254-14267 from: 1 to: 158

```
63 SerIeuSerIeuArgSerA 69
|||||
7  TCCCTCTCTGCGCTGCGCA 27
```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-673-476-266

seq_documentation_block:

```
; Sequence 266, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESEN-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/09/673,476
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO: 266
; LENGTH: 217
; TYPE: DNA
```



```
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (139)..(140)
; OTHER INFORMATION: a, t, c or g
US-09-673-476-266
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-673-476-266 ..
```

```
Align seg 1/1 to: US-09-673-476-266 from: 1 to: 217
```

```
63 SerleusSerleuArgSerAla 69
|||||
10 TCCTTATCGCTCCGCTCTGCA 30
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8128
```

seq_documentation_block:

```
; Sequence 8128, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; SOFTWARE: PERL Program
; SEQ ID NO 8128
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00121720
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: 60, 178
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-800C-8128
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-539-800C-8128 ..
```

```
Align seg 1/1 to: US-09-539-800C-8128 from: 1 to: 229
```

```
63 SerleusSerleuArgSerAla 69
|||||
92 TCACCTTCGCTGCTGCTGCT 112
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-23834
```

seq_documentation_block:

```
; Sequence 23834, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 23834
; LENGTH: 234
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700962842H1
US-09-975-254-23834
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-975-254-23834/rev ..
```

```
Align seg 1/1 to reverse of: US-09-975-254-23834 from: 1 to: 234
```

```
63 SerleusSerleuArgSerAla 69
|||||
173 TCTCTAAGCTACGCTGCTGCT 153
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-13594
```

seq_documentation_block:

```
; Sequence 13594, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
```



```
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 13594
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00527775
; NAME/KEY: unsure
; LOCATION: 206
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-13594

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-331D-13594 ..
Align seg 1/1 to: US-09-539-331D-13594 from: 1 to: 239

153 LeuAsn11eAlaProAlaGlu 159
|||||
119 CTCACATTGTCACGACGAG 139

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11353
seq_documentation_block:
; Sequence 11353, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
```

```
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 11353
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00067358
; NAME/KEY: unsure
; LOCATION: 7, 93, 186, 204
; OTHER INFORMATION: a, t, c, g, or other
```


US-09-540-210B-11353

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-09-540-210B-11353

Align seg 1/1 to: US-09-540-210B-11353 from: 1 to: 242

217 LysValLysArgGlnIlePhe 223

143 AAGTAAAGACAGCATATT 163

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-9255

seq_documentation_block:

```
; Sequence 9255, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO: 9255
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00523285
US-09-539-331D-9255
```

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-09-539-331D-9255/rev

Align seg 1/1 to reverse of: US-09-539-331D-9255 from: 1 to: 253

66 LeuArgSerAlaHisLeuAla 72

119 TTACGGCTCTCACACCTGGCC 99

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235

seq_documentation_block:

```
; Sequence 8235, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
```

CURRENT FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 08/521,383

PRIOR FILING DATE: August 16, 1995

PRIOR APPLICATION NUMBER: 08/271,217

PRIOR FILING DATE: June 27, 1994

PRIOR APPLICATION NUMBER: 08/334,881

PRIOR FILING DATE: November 4, 1994

PRIOR APPLICATION NUMBER: 08/943,978

PRIOR FILING DATE: October 3, 1997

PRIOR APPLICATION NUMBER: 60/028,732

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/943,979

PRIOR FILING DATE: October 4, 1997

PRIOR APPLICATION NUMBER: 60/027,782

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/993,774

PRIOR FILING DATE: December 18, 1997

PRIOR APPLICATION NUMBER: 60/034,975

PRIOR FILING DATE: December 20, 1996

PRIOR APPLICATION NUMBER: 09/250,003

PRIOR FILING DATE: February 10, 1999

PRIOR APPLICATION NUMBER: 60/074,364

PRIOR FILING DATE: February 12, 1998

PRIOR APPLICATION NUMBER: 09/452,747

PRIOR FILING DATE: December 1, 1999

PRIOR APPLICATION NUMBER: 60/111,910

PRIOR FILING DATE: December 10, 1998

NUMBER OF SEQ ID NOS: 19698

SOFTWARE: PERL Program

SEQ ID NO: 8235

LENGTH: 272

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu00266046

US-09-539-800C-8235

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-09-539-800C-8235/rev

Align seg 1/1 to reverse of: US-09-539-800C-8235 from: 1 to: 272

7 ArgAlaSpSerArgPropio 13

122 AGGGCTGCTCTCGGCCCCC 102

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360

seq_documentation_block:

```
; Sequence 15360, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
```


PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/005,526
PRIOR FILING DATE: September 29, 1995
PRIOR APPLICATION NUMBER: 08/824,029
PRIOR FILING DATE: March 25, 1997
PRIOR APPLICATION NUMBER: 60/014,010
PRIOR FILING DATE: March 25, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/903,555
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/862,178
PRIOR FILING DATE: May 22, 1997
PRIOR APPLICATION NUMBER: 60/018,217
PRIOR FILING DATE: May 23, 1996
PRIOR APPLICATION NUMBER: 08/881,589
PRIOR FILING DATE: June 24, 1997
PRIOR APPLICATION NUMBER: 60/021,275
PRIOR FILING DATE: June 25, 1996
PRIOR APPLICATION NUMBER: 08/903,802
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/905,881
PRIOR FILING DATE: August 1, 1997
PRIOR APPLICATION NUMBER: 60/025,204
PRIOR FILING DATE: August 1, 1996
PRIOR APPLICATION NUMBER: 08/903,471
PRIOR FILING DATE: July 30, 1997
PRIOR APPLICATION NUMBER: 60/025,478
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/903,556
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/025,217
PRIOR FILING DATE: August 22, 1996
PRIOR APPLICATION NUMBER: 08/937,142
PRIOR FILING DATE: September 23, 1997
PRIOR APPLICATION NUMBER: 60/026,598
PRIOR FILING DATE: September 24, 1996
PRIOR APPLICATION NUMBER: 08/960,746
PRIOR FILING DATE: October 29, 1997
PRIOR APPLICATION NUMBER: 60/030,144
PRIOR FILING DATE: October 30, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/755,524
PRIOR FILING DATE: November 22, 1996
PRIOR APPLICATION NUMBER: 60/007,495
PRIOR FILING DATE: November 22, 1995
PRIOR APPLICATION NUMBER: 09/021,031
PRIOR FILING DATE: February 10, 1998
PRIOR APPLICATION NUMBER: 60/039,325
PRIOR FILING DATE: February 13, 1997
PRIOR APPLICATION NUMBER: 09/035,172
PRIOR FILING DATE: March 4, 1998
PRIOR APPLICATION NUMBER: 60/040,431
PRIOR FILING DATE: March 5, 1997
PRIOR APPLICATION NUMBER: 09/041,894
PRIOR FILING DATE: March 12, 1998
PRIOR APPLICATION NUMBER: 60/040,199
PRIOR FILING DATE: March 14, 1997
PRIOR APPLICATION NUMBER: 09/050,817
PRIOR FILING DATE: March 30, 1998
PRIOR APPLICATION NUMBER: 60/043,792
PRIOR FILING DATE: April 11, 1997
PRIOR APPLICATION NUMBER: 09/074,999
PRIOR FILING DATE: May 8, 1998

PRIOR APPLICATION NUMBER: 60/048,431
PRIOR FILING DATE: May 29, 1997
PRIOR APPLICATION NUMBER: 09/107,592
PRIOR FILING DATE: June 30, 1998
PRIOR APPLICATION NUMBER: 60/052,751
PRIOR FILING DATE: July 1, 1997
PRIOR APPLICATION NUMBER: 09/094,079
PRIOR FILING DATE: June 9, 1998
PRIOR APPLICATION NUMBER: 60/049,975
PRIOR FILING DATE: June 13, 1997
NUMBER OF SEQ ID NOS: 35654
SOFTWARE: PERL Program
SEQ ID NO 15360
LENGTH: 279
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu00456393
US-09-540-210B-15360

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-540-210B-15360/rev ..
Align seg 1/1 to reverse of: US-09-540-210B-15360 from: 1 to: 279

62 ThresleuSerleuArgSer 68
.....
45 ACCAGCCTCTCCCTGAGGAGC 25

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-20652

seq_documentation_block:
Sequence 20652, Application US/09540210B
GENERAL INFORMATION:
APPLICANT: Sellhamer, Jeffrey J.
APPLICANT: Deleagane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Muliahy, Sara J.
TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
FILE REFERENCE: PD-1037 CIP
CURRENT APPLICATION NUMBER: US/09/540,210B
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 08/972,899
PRIOR FILING DATE: November 18, 1997
PRIOR APPLICATION NUMBER: 08/395,244
PRIOR FILING DATE: February 27, 1995
PRIOR APPLICATION NUMBER: 08/722,922
PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/005,526
PRIOR FILING DATE: September 29, 1995
PRIOR APPLICATION NUMBER: 08/824,029
PRIOR FILING DATE: March 25, 1997
PRIOR APPLICATION NUMBER: 60/014,010
PRIOR FILING DATE: March 25, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/903,555
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/862,178
PRIOR FILING DATE: May 22, 1997


```

; PRIOR APPLICATION NUMBER: 60/018, 217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881, 589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021, 275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903, 802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023, 308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905, 881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025, 204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903, 471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025, 478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903, 556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025, 217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937, 142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026, 598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960, 746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030, 144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826, 847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015, 533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755, 524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007, 495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021, 031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039, 325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035, 172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040, 431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041, 894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040, 199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050, 817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043, 792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074, 999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048, 431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107, 592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052, 751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094, 079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049, 975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 20652
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00094015
; NAME/KEY: unsure
; LOCATION: 189-190, 229
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-210B-20652
```

```

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
```

```

alignment_block:
US-09-528-682-1 x US-09-540-210B-20652/rev ..
```

```

Align seg 1/1 to reverse of: US-09-540-210B-20652 from: 1 to: 282
```

```

66 leuArgSerAlaHisLeuAla 72
|||||
75 TTACGGCTGTCACACCTGGCC 55
```

```

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303
```

```

seq_documentation_block:
; Sequence 28303, Application US/09539331D
```

```

; GENERAL INFORMATION:
```

```

; APPLICANT: Seilhamer, Jeffrey J.
```

```

; APPLICANT: Delegeane, Angelo M.
```

```

; APPLICANT: Stuart, Susan G.
```

```

; APPLICANT: Stuve, Laura L.
```

```

; APPLICANT: Mullahy, Sara J.
```

```

; APPLICANT: Naughton, Rebecca E.
```

```

; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
```

```

; FILE REFERENCE: PD-1022 CIP
```

```

; CURRENT APPLICATION NUMBER: US/09/539, 331D
```

```

; PRIOR FILING DATE: 2000-03-30
```

```

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
```

```

; NUMBER OF SEQ ID NOS: 40961
```

```

; SOFTWARE: PERL Program
```

```

; SEQ ID NO 28303
```

```

; LENGTH: 312
```

```

; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
```

```

; FEATURE:
```

```

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00114685
```

```

; US-09-539-331D-28303
```

```

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000
```

```

alignment_block:
US-09-528-682-1 x US-09-539-331D-28303 ..
```

```

Align seg 1/1 to: US-09-539-331D-28303 from: 1 to: 312
```

```

63 SerLeuSerLeuArgSerAla 69
|||||
31 AGCCTGAGTCTGAGAGTCCA 51
```

```

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-22417
```

```

seq_documentation_block:
; Sequence 22417, Application US/09539331D
```

```

; GENERAL INFORMATION:
```

```

; APPLICANT: Seilhamer, Jeffrey J.
```

```

; APPLICANT: Delegeane, Angelo M.
```

```

; APPLICANT: Stuart, Susan G.
```

```

; APPLICANT: Stuve, Laura L.
```



```
; APPLICANT: Mullaby, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539, 331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 22417
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00325793
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 45, 51, 64, 78, 93, 108, 144, 150, 168, 216, 234, 249
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-22417

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-331D-22417/rev ..

Align seg 1/1 to reverse of: US-09-539-331D-22417 from: 1 to: 318

66 leuargSerleuargSerala 72
|||||
206 CTAGGCTCGCGCAGCTTGCA 186

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-10421-160

seq_documentation_block:
; Sequence 160, Application PC/TUS0210421
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Siqing
; APPLICANT: Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF CANCER
; FILE REFERENCE: 210121.565PC
; CURRENT APPLICATION NUMBER: PCT/US02/10421
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2959
; SEQ ID NO 160
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-10421-160

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US02-10421-160 ..

Align seg 1/1 to: PCT-US02-10421-160 from: 1 to: 323

63 SerleuargSerleuargSerala 69
|||||
71 TCTCTTCTCTCTCAGAGTGCT 91
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-15587

seq_documentation_block:
; Sequence 15587, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Afterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomeili, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raist, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Rei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721, 544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515, 128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034, 341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15587
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-544-15587

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-721-544-15587/rev ..

Align seg 1/1 to reverse of: US-09-721-544-15587 from: 1 to: 323

61 SerThrSerleuargSerala 67
|||||
263 TCAACATCCTACTGCTCTTGAGG 243

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-699-160

seq_documentation_block:
; Sequence 160, Application US/10112699
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Siqing
; APPLICANT: Chaitanya S.
```



```
; APPLICANT: Gaiger, Alexander
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.565
; CURRENT APPLICATION NUMBER: US/10/112.699
; NUMBER OF SEQ ID NOS: 2959
; SEQ ID NO 160
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-112-699-160

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-112-699-160 ..

Align seg 1/1 to: US-10-112-699-160 from: 1 to: 323

63 SerLeuSerLeuArgSerAla 69
|||||
71 TCCTCTCTCTCAGAGTGTCT 91

seq_name: /cgn2/6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-15776

seq_documentation_block:
; Sequence 15776, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; SOFTWARE: PERL Program
; SEQ ID NO 15776
; LENGTH: 324
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01286321
; US-09-539-800C-15776

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-800C-15776/rev ..

Align seg 1/1 to reverse of: US-09-539-800C-15776 from: 1 to: 324

112 GluValSerAlaLeuGlycyl 118
|||||
85 GAGTAAGTGCACCTTGAGGCG 65

seq_name: /cgn2/6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800B-4227

seq_documentation_block:
; Sequence 4227, Application US/09539806B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539,806B
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/706,766
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/004,676
; PRIOR FILING DATE: October 2, 1995
; PRIOR APPLICATION NUMBER: 08/749,515
; PRIOR FILING DATE: November 15, 1996
; PRIOR APPLICATION NUMBER: 60/006,810
; PRIOR FILING DATE: November 15, 1995
; PRIOR APPLICATION NUMBER: 08/822,285
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/013,696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/951,197
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027,249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/826,438
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/016,145
; PRIOR FILING DATE: April 12, 1996
; PRIOR APPLICATION NUMBER: 60/013,696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/839,389
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 60/015,312
; PRIOR FILING DATE: April 12, 1996
; PRIOR APPLICATION NUMBER: 08/951,198
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027,249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/839,968
; PRIOR FILING DATE: April 24, 1997
; PRIOR APPLICATION NUMBER: 60/016,150
; PRIOR FILING DATE: April 24, 1996
; PRIOR APPLICATION NUMBER: 08/846,104
; PRIOR FILING DATE: April 25, 1997
```



```

? PRIOR APPLICATION NUMBER: 60/016,838
? PRIOR FILING DATE: April 26, 1996
? PRIOR APPLICATION NUMBER: 08/999,886
? PRIOR FILING DATE: May 13, 1997
? PRIOR APPLICATION NUMBER: 60/017,555
? PRIOR FILING DATE: May 13, 1996
? PRIOR APPLICATION NUMBER: 60/031,667
? PRIOR FILING DATE: November 11, 1996
? PRIOR APPLICATION NUMBER: 08/901,902
? PRIOR FILING DATE: July 24, 1997
? PRIOR APPLICATION NUMBER: 60/023,379
? PRIOR FILING DATE: July 26, 1996
? PRIOR APPLICATION NUMBER: 08/918,181
? PRIOR FILING DATE: August 27, 1997
? PRIOR APPLICATION NUMBER: 60/023,236
? PRIOR FILING DATE: August 28, 1996
? PRIOR APPLICATION NUMBER: 08/940,864
? PRIOR FILING DATE: September 29, 1997
? PRIOR APPLICATION NUMBER: 60/027,236
? PRIOR FILING DATE: September 30, 1996
? PRIOR APPLICATION NUMBER: 08/956,502
? PRIOR FILING DATE: October 22, 1997
? PRIOR APPLICATION NUMBER: 60/029,083
? PRIOR FILING DATE: October 23, 1996
? PRIOR APPLICATION NUMBER: 08/993,402
? PRIOR FILING DATE: December 18, 1997
? PRIOR APPLICATION NUMBER: 60/033,647
? PRIOR FILING DATE: December 19, 1996
? PRIOR APPLICATION NUMBER: 09/022,355
? PRIOR FILING DATE: February 11, 1998
? PRIOR APPLICATION NUMBER: 60/044,847
? PRIOR FILING DATE: February 13, 1997
? PRIOR APPLICATION NUMBER: 09/045,574
? PRIOR FILING DATE: March 20, 1998
? PRIOR APPLICATION NUMBER: 60/041,275
? PRIOR FILING DATE: March 21, 1997
? PRIOR APPLICATION NUMBER: 09/042,629
? PRIOR FILING DATE: March 16, 1998
? PRIOR APPLICATION NUMBER: 60/043,613
? PRIOR FILING DATE: March 18, 1997
? PRIOR APPLICATION NUMBER: 09/057,988
? PRIOR FILING DATE: April 9, 1998
? PRIOR APPLICATION NUMBER: 60/043,256
? PRIOR FILING DATE: April 16, 1997
? PRIOR APPLICATION NUMBER: 09/114,041
? PRIOR FILING DATE: July 10, 1998
? PRIOR APPLICATION NUMBER: 60/052,257
? PRIOR FILING DATE: July 10, 1997
? PRIOR APPLICATION NUMBER: 09/145,340
? PRIOR FILING DATE: September 1, 1998
? PRIOR APPLICATION NUMBER: 60/058,922
? PRIOR FILING DATE: September 11, 1997
? NUMBER OF SEQ ID NOS: 48372
? SOFTWARE: PERL Program
? SEQ ID NO 4227
? LENGTH: 349
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Incyte ID No: hu00577412
? FEATURE:
? NAME/KEY: unsure
? LOCATION: 37, 56, 60, 62, 66, 69
? OTHER INFORMATION: a, t, c, g, or other
? US-09-539-806B-4227

```

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 7
  Gaps: 0
  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x US-09-539-806B-4227
  Align seg 1/1 to: US-09-539-806B-4227 from: 1 to: 349

```

```

114 SerAlaLeuGlyIlePro 120
|||||
264 TCTGCATTAAGGGGCGATCCCA 284

```

```
seq_name: /cgn2_6/prodata/1/pna/us09_NEW_COMB.seq:US-09-721-544-19743
```

```

seq_documentation_block:
? Sequence 19743, Application US/09721544
? GENERAL INFORMATION:
? APPLICANT: Arterburn, Matthew
? APPLICANT: Asghari, Vida
? APPLICANT: Damavandi, Simin
? APPLICANT: Dickson, Mark
? APPLICANT: Drake, Jim
? APPLICANT: Drmanac, Radoje
? APPLICANT: Engleman, Carrie
? APPLICANT: Faulkner, Brandy
? APPLICANT: Garcia, Veronica
? APPLICANT: Giedt, Gretchen
? APPLICANT: Hunter, Kelly
? APPLICANT: Jensen, Aaron
? APPLICANT: Jones, Lee
? APPLICANT: Kita, David
? APPLICANT: Labat, Ivan
? APPLICANT: Laroza, Mimi
? APPLICANT: Lomelli, Michelle
? APPLICANT: Nguyen, Phuong
? APPLICANT: Nogra, Margie
? APPLICANT: Palencia, Servando
? APPLICANT: Ralsi, Fariba
? APPLICANT: Smith, Benjamin
? APPLICANT: Tkach, Joe
? APPLICANT: Tran, Lien
? APPLICANT: Verna, Ron
? APPLICANT: Yang, Fel
? APPLICANT: Yim, Kenneth
? TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
? FILE REFERENCE: 728CIP
? CURRENT APPLICATION NUMBER: US/09/721,544
? PRIOR FILING DATE: 2000-11-21
? PRIOR APPLICATION NUMBER: 09/515,128
? PRIOR FILING DATE: 1998-02-13
? PRIOR APPLICATION NUMBER: 09/034,341
? PRIOR FILING DATE: 1998-02-13
? NUMBER OF SEQ ID NOS: 24489
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 19743
? LENGTH: 356
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-721-544-19743

```

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 7
  Gaps: 0
  Percent Identity: 100.000

```

```
alignment_block:
  US-09-528-682-1 x US-09-721-544-19743
```

```
Align seg 1/1 to: US-09-721-544-19743 from: 1 to: 356
```

```

61 SerThrSerLeuSerLeuArg 67
|||||
27 AGCACTTCATCAGTTAAGA 47

```


seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-154-2094

seq_documentation_block:

; Sequence 2094, Application US/10011154
; GENERAL INFORMATION:
; APPLICANT: Dmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark C.
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 780CIP
; CURRENT APPLICATION NUMBER: US/10/011,154
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/524,038
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/404,284
; NUMBER OF SEQ ID NOS: 4670
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 2094
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-154-2094

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-10-011-154-2094/rev ..

Align seg 1/1 to reverse of: US-10-011-154-2094 from: 1 to: 410

65 SerleuArgSerAlaHisLeu 71
|||||
353 TCCCTTCGCTCGTCATCTC 333

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-3305

seq_documentation_block:

; Sequence 3305, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3305
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-3305

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-3305/rev ..

Align seg 1/1 to reverse of: US-10-027-632-3305 from: 1 to: 435

61 SerThrSerLeuArg 67
|||||
103 TCAACTCGTGGAGCCACGA 83

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-154-2817

seq_documentation_block:

; Sequence 2817, Application US/10011154
; GENERAL INFORMATION:
; APPLICANT: Dmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark C.
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 780CIP
; CURRENT APPLICATION NUMBER: US/10/011,154
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/524,038
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/404,284
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 4670
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 2817
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-154-2817

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-10-011-154-2817 ..

Align seg 1/1 to: US-10-011-154-2817 from: 1 to: 436

65 SerLeuArgSerAlaHisLeu 71
|||||
283 TCCCTTCGCTCGTCATCTC 303

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-89412

seq_documentation_block:

; Sequence 89412, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676


```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89412
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89412
```

```
alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-89412 ..

Align seg 1/1 to: US-10-027-632-89412 from: 1 to: 448

```
60 ValSerThrSerLeuSerIeu 66
|||||
335 GTGCTACATCTCTCTCTC 355
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-89413

seq_documentation_block:

```
; Sequence 89413, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89413
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89413
```

```
alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-89413 ..

Align seg 1/1 to: US-10-027-632-89413 from: 1 to: 448

```
60 ValSerThrSerLeuSerIeu 66
|||||
335 GTGCTACATCTCTCTCTC 355
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-144-702-194

seq_documentation_block:

```
; Sequence 194, Application US/10144702
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PMO10CIN
; CURRENT APPLICATION NUMBER: US/10/144,702
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/758,460
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 940
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 194
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (456)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (457)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-144-702-194
```

```
alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-144-702-194/rev ..

Align seg 1/1 to reverse of: US-10-144-702-194 from: 1 to: 457

```
62 ThrSerLeuArgSer 68
|||||
99 ACTTCCCTGAGCCACAGTCC 79
```

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-2302

seq_documentation_block:

```
; Sequence 2302, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
```



```

1  APPLICANT: Garcia, Veronica
2  APPLICANT: Giedt, Gretchen
3  APPLICANT: Hunter, Kelly
4  APPLICANT: Jessen, Aaron
5  APPLICANT: Jones, Lee
6  APPLICANT: Kita, David
7  APPLICANT: Labat, Ivan
8  APPLICANT: Laroya, Mimi
9  APPLICANT: Lomelli, Michelle
10 APPLICANT: Nguyen, Phuong
11 APPLICANT: Nogra, Margie
12 APPLICANT: Palencia, Servando
13 APPLICANT: Raisi, Fariba
14 APPLICANT: Smith, Benjamin
15 APPLICANT: Tkach, Joe
16 APPLICANT: Tran, Lien
17 APPLICANT: Verna, Ron
18 APPLICANT: Yang, Fei
19 APPLICANT: Yim, Kenneth
20 TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
21 FROM CDNA LIBRARY OF FETAL LIVER-SPLEEN
22 FILE REFERENCE: 728CIP
23 CURRENT APPLICATION NUMBER: US/09/721,544
24 CURRENT FILING DATE: 2000-11-21
25 PRIOR APPLICATION NUMBER: 09/515,128
26 PRIOR FILING DATE: 1998-02-13
27 PRIOR APPLICATION NUMBER: 09/034,341
28 PRIOR FILING DATE: 1998-02-13
29 NUMBER OF SEQ ID NOS: 24489
30 SOFTWARE: FastSeq for Windows Version 3.0
31 SEQ ID NO 2302
32 LENGTH: 470
33 TYPE: DNA
34 ORGANISM: Homo sapiens
35 US-09-721-544-2302

```

```
alignment_scores:      7
      quality:      7.00
      ratio:      1.000
Percent Similarity: 100.000
Percent Identity: 100.0000
```

alignment_block:
US-09-528-682-1 x US-09-721-544-2302 .

Align seg 1/1 to: US-09-721-544-2302 from: 1 to: 470

187 CysGlyAsnSerSerArgThr 193
 259 TGTGGCAATTCCTCAAGACC 279

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-704-302A-158

```

seq_documentation_block:
  Sequence 158, Application US/09704302A
  GENERAL INFORMATION:
    APPLICANT: Strabala, Timothy
    APPLICANT: Nieuwenhuizen, Nicolaas
    TITLE OF INVENTION: Compositions Isolated from Plant Cells
    TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
    FILE REFERENCE: 11000.1020c1u
    CURRENT APPLICATION NUMBER: US/09/704, 302A
    CURRENT FILING DATE: 2002-05-10
    NUMBER OF SEQ ID NOS: 1402
    SOFTWARE: FastSeq for Windows Version 4.0
    SEQ ID NO 158
    LENGTH: 471
    TYPE: DNA
    ORGANISM: Eucalyptus grandis
  US-09-704-302A-158

```

alignment_scores:

quality:	7.00	length:	7
ratio:	1.000	gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block: NS-09-528-682-1 x NS-09-704-302A-158/rev

Align seg 1/1 to reverse of: US-09-704-302A-158 from: 1 to: 471

71 LeuAlaGlyInserIleLeu 77
 |||||
 361 CTTGCCGGACAATCCATTCTC 341

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-355

```

seq.documentation_block:
? Sequence 355, Application US/10002623
? GENERAL INFORMATION:
?   APPLICANT: OEFFNER, PETER J.
?   APPLICANT: UNDERHILL, PETER A.
?   TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
?   TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMANS
?   TITLE OF INVENTION: POPULATIONS
?   FILE REFERENCE: STAN-212
?   CURRENT APPLICATION NUMBER: US/10/002,623
?   CURRENT FILING DATE: 2001-11-01
?   PRIOR APPLICATION NUMBER: US 60/245,355
?   PRIOR FILING DATE: 2000-11-01
?   NUMBER OF SEQ ID NOS: 952
?   SOFTWARE: FastSeq for Windows Version 4.0
?   SEQ ID NO 355
?   LENGTH: 478
?   TYPE: DNA
?   ORGANISM: Homo Sapiens
?   S-10-002-623-355

```

```
alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Gaps: 0
Percent Similarity: 100.000
Percent Identity: 100.000
```

alignment_block:
US-09-528-682-1 x US-10-002-623-355/rev . .

Align seg 1/1 to reverse of: US-10-002-623-355 from: 1 to: 478

```

110 GluGluValSerAlaLeu 116
      |||||
323 GAGCAGGAGGTCTCAGCTCTA 303

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-86877

```

seq.documentation.block:
  Sequence 86877, Application US/10027632
  GENERAL INFORMATION:
    APPLICANT: Wang, David G.
    TITLE OF INVENTION: Identification and Mapping of Single Nucleotides
    TITLE OF INVENTION: Polymorphisms in the Human Genome
    FILE REFERENCE: 108827.129
    CURRENT APPLICATION NUMBER: US/10/027,632
    PRIOR APPLICATION NUMBER: US 60/218,006
    PRIOR FILING DATE: 2000-07-12
    PRIOR APPLICATION NUMBER: US 60/198,676
    PRIOR FILING DATE: 2000-04-20
    PRIOR APPLICATION NUMBER: US 60/199,483
    PRIOR FILING DATE: 2000-03-29
    PRIOR APPLICATION NUMBER: US 60/185,218
    PRIOR FILING DATE: 2000-02-24
    PRIOR APPLICATION NUMBER: US 60/167,363
    PRIOR FILING DATE: 1999-11-23
    PRIOR APPLICATION NUMBER: US 60/156,358

```

alignment_scores:


```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86877
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86877

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-86877  ..

Align seg 1/1 to: US-10-027-632-86877 from: 1 to: 478

144 GUTYrArgAspArgTyrTyr 150
|||||
224 GAATATAGAGACAGTATTAT 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-86878

seq_documentation_block:
; Sequence 86878, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86878
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86878

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-86878  ..

Align seg 1/1 to: US-10-027-632-86878 from: 1 to: 478

144 GUTYrArgAspArgTyrTyr 150
|||||
224 GAATATAGAGACAGTATTAT 244
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-3385

seq_documentation_block:
; Sequence 3385, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radcoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jensen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomeili, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smich, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Rei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 728CIP
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US/09/721,544
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3385
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(488)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-544-3385

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-721-544-3385/rev  ..

Align seg 1/1 to reverse of: US-09-721-544-3385 from: 1 to: 488

62 ThrSerLeuSerLeuArgSer 68
|||||
265 ACATCATTAAGCCTCGAACT 245

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-68521

seq_documentation_block:
; Sequence 68521, Application US/10027632
```



```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68521
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68521

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-68521/rev ..

Align seg 1/1 to reverse of: US-10-027-632-68521 from: 1 to: 493

208 ThrIleTyLeuArgGluTyr 214
|||||
403 ACTATCTATTTAAGAGAGTAC 383

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-295156
seq_documentation_block:
; Sequence 295156, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 295156
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-295156
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-295156

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-295156/rev ..

Align seg 1/1 to reverse of: US-10-027-632-295156 from: 1 to: 493

208 ThrIleTyLeuArgGluTyr 214
|||||
403 ACTATCTATTTAAGAGAGTAC 383

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-60190
seq_documentation_block:
; Sequence 60190, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60190
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-60190

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-60190/rev ..

Align seg 1/1 to reverse of: US-10-027-632-60190 from: 1 to: 502

58 GLTYrVAlserThrSerLeu 64
|||||
485 GGGTATGTTCTACTTCACATC 465

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-298567
seq_documentation_block:
; Sequence 298567, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```



```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 298567
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-298567

```

```

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x US-10-027-632-298567/rev ..

```

```

Align seg 1/1 to reverse of: US-10-027-632-298567 from: 1 to: 502

```

```

58 GlyTyrValSerThrSerLeu 64
|||||
485 GGGTATGTTCTACTCCTC 465

```


